

Un processus pour la sélection de stratégie de gestion de l'information dans les systèmes de confiance

N. Cointe, A. Ghorbani, C. Chorus

Department of Engineering Systems and Services,
Faculty of Technology, Policy and Management,
Delft University of Technology, Pays-Bas

nicolas.cointe@tudelft.nl

a.ghorbani@tudelft.nl

c.g.chorus@tudelft.nl

Résumé

Dans certains systèmes multi-agents, en particulier des systèmes ouverts et distribués, la coopération peut nécessiter au préalable d'observer et évaluer le comportement des autres. Pour ce faire, plusieurs techniques de reconnaissance d'objectifs et de buts ont déjà été présentées dans la littérature. Nous introduisons ici le concept de gestion de l'information, permettant à l'agent autonome artificiel de mesurer et connaître la quantité d'information donnée à un observateur pour en tenir compte dans sa prise de décisions, et montrons comment cela peut impacter ses interactions sociales. Un ensemble d'expérimentations employant des agents BDI illustre enfin la sensibilité de ces interactions sociales aux stratégies de gestion de l'information utilisées.

Mots-clés

Confiance, auto-organisation, confidentialité.

Abstract

In some multiagent systems, more particularly open and distributed systems, coalitions formation requires to observe and evaluate the others' behavior before initiating any cooperation. In that way, the literature already proposes some goals and plans recognition techniques. We introduce here the concept of Information regulation, allowing artificial autonomous agents to measure – and be aware of – the quantity of information given to an observer and consider it to make decisions. A set of experiments illustrates how the various information regulation strategies may impact social interactions of BDI agents.

Keywords

Trust, self-organization, privacy.

1 Introduction

Dans des jeux coopératifs, trouver des partenaires susceptibles d'être intéressés par une coopération et la construction d'une coalition ou d'une organisation est une tâche complexe et sensible, lorsqu'une telle interaction peut influencer les bénéfices et pertes des agents ou (non exclusif) les exposer à des sanctions sociales telles qu'une exclusion.

De tels cas de figures peuvent se présenter par exemple dans des situations présentant une dimension morale [6] et où des actions, bien qu'exécutables et permises, peuvent rendre leurs auteurs désirables ou non au sein d'un collectif. Notons que nous traitons ici de cas où les agents ne disposent d'aucune preuve sur les intentions des autres agents mais emploient des techniques proposées dans la littérature afin d'évaluer avec une certitude suffisante, les objectifs d'un agent.

Sélectionner des partenaires semblant partager des objectifs similaires peut se faire, de manière classique, à l'aide de modèles de confiance [21] afin de laisser le système s'auto-organiser. L'avantage d'un tel modèle est qu'il est construit de manière décentralisée, par les agents, en se fondant sur leur expérience individuelle – tirant profit d'observations, connaissances et raisonnement propres à l'agent – et éventuellement sur des expériences transmises par un système de réputation [26].

Afin de pouvoir raisonner sur les objectifs des autres agents, la littérature propose déjà un ensemble de techniques diverses, que nous regroupons dans cet article au sein du concept plus général de *reconnaissance d'objectifs*, permettant d'inférer, à partir d'observations, le but le plus vraisemblable compte tenu du comportement d'un agent observé. Les méthodes de reconnaissance d'objectif ont démontré ces deux dernières décennies leur intérêt dans divers domaines applicatifs nécessitant une coopération entre agents autonomes et humains, tels que l'apprentissage de routines comportementales pour la surveillance de personnes âgées à domicile [14], ou bien pour la coopération d'agents autonomes entre eux, par exemple dans le cas d'un groupe de robots devant coordonner leurs déplacements en minimisant les communications [4]. La reconnaissance d'objectifs est également utile dans un cadre compétitif, afin d'anticiper et éventuellement contrer les buts d'un adversaire. Par exemple, leur usage a montré sa pertinence pour la détection de comportements humains dangereux dans une foule [22], ou la déduction de la stratégie employée par un adversaire dans des jeux vidéos [15] ou des compétitions de football robotiques [16].

Cependant, tenir compte de l'observation de leur comportement, et de la présence d'un système de confiance

et de réputation peut permettre à des agents –ou groupes d’agents– d’influencer l’opinion des autres agents du système afin de les tromper sur leurs propres intentions [2] ou seulement maintenir un certain niveau d’incertitude chez l’observateur [7]. Dans le cadre de cet article, nous exposons le cas d’agents cherchant à maximiser ou à minimiser l’information donnée au travers de leur comportement à des observateurs afin de clarifier ou dissimuler leurs objectifs dans le but d’influencer leurs chances de rejoindre un collectif d’agents, ou de minimiser les risques de sanctions.

Le concept de *gestion d’information* propose d’évaluer la quantité d’information pouvant être révélée par un comportement (exécuté ou anticipé) à un potentiel observateur. Cette mesure de quantité d’information transmise permet alors d’employer une stratégie de minimisation de l’information, nous parlerons alors de *dissimulation*, ou au contraire de sa maximisation, que nous désignons par le terme de *transparence*.

Cet article propose d’explorer divers aspects des stratégies de gestion d’information au travers d’un modèle dans lequel chacun cherche à identifier les buts des autres et gagner la confiance de ceux qu’il estime poursuivre un objectif similaire ou compatible afin de construire un collectif.

En Section 2, nous présentons des travaux de la littérature présentant diverses contributions en matière de reconnaissance d’objectif, de gestion d’information et de mécanismes de construction de confiance. Puis en Section 3 nous présentons notre proposition afin de décrire le processus selon lequel les agents adoptent une stratégie de régulation de l’information en fonction du contexte. La Section 4 illustre des résultats expérimentaux obtenus à la suite de l’implémentation de ce modèle, et son usage dans un scénario montrant des agents dotés de points de vue et objectifs divers. Enfin, en Section 5, nous présentons nos conclusions sur l’intérêt de cette contribution et présentons nos intentions pour la poursuite de nos travaux.

2 État de l’art

Dans cette section, nous exposons des travaux illustrant plus en détail les concepts fondamentaux sur lesquels repose notre proposition.

En premier lieu, nous définissons le comportement b_{ag_i} d’un agent $ag_i \in Ag$ comme un ensemble ordonné d’actions :

$$b_{ag_i} = \{a_{ag_i,t_0}, \dots, a'_{ag_i,t'}\}$$

avec $a_{ag_i,t_0}, \dots, a'_{ag_i,t'}$ une séquence d’actions exécutée par l’agent ag_i entre les instants t_0 et t' . Nous désignons dans le cadre de cet article sous l’intitulé de *comportement observé*, un ensemble d’actions observées ayant déjà été exécutées par l’agent ag_i tel qu’à l’instant courant t , $t' \leq t$. À l’inverse, un *comportement hypothétique* inclut au moins une action hypothétique, c’est-à-dire dont l’exécution n’a pas été observée. Cette absence d’observation peu s’expliquer par la nature anticipée d’un comportement, dans le cas où $t' > t$, soit qu’un agent cherche à évaluer un plan qu’il a l’intention d’exécuter, soit qu’il en-

visage un comportement futur possible d’un autre agent. Cela peut aussi permettre, dans des systèmes partiellement observables, d’envisager des actions qui n’ont pas été observées bien qu’elles puissent avoir été effectuées.

La sous-section 2.1 présente premièrement plusieurs aspects de la modélisation de la confiance dans les sociétés artificielles. En sous-section 2.2, nous présentons un ensemble non-exhaustif de techniques de reconnaissance d’objectifs, afin d’illustrer leur diversité et proposer en conséquence une définition générique de *fonction de reconnaissance d’objectif* (GRF) englobant l’ensemble de ces propositions. Enfin, la sous-section 2.3 montre comment une fonction de reconnaissance d’objectif peut être utilisée dans un processus décisionnel afin de mettre en œuvre une stratégie de reconnaissance d’objectif.

2.1 Modèles de confiance

Dans un système multi-agent ouvert, décentralisé et hétérogène, les agents peuvent chercher à atteindre divers objectifs, avoir des perceptions et actions diverses, voire des architectures entièrement différentes. Afin de permettre aux agents d’évaluer dynamiquement les autres agents selon leurs propres critères et architecture, afin de sélectionner de potentiels collaborateurs, une approche classique consiste à implémenter un processus de confiance [21]. Dans un tel modèle chaque agent évalue les autres en fonction de l’observation de leur comportement et des évaluations successives. L’état mental, construit par ces évaluations est appelé *image* et prend généralement une valeur, discrète ou continue qui, lorsqu’elle est considérée suffisante, peut entraîner l’établissement d’une confiance [26]. L’une des définitions les plus consensuelles [18, 21] de la confiance dans le domaine des systèmes multi-agents, semble être celle proposée par Gambetta [9] la définissant comme : “un niveau particulier de la probabilité subjective qu’un agent –ou groupe d’agent– puisse effectuer une action”. Ici, plus formellement, nous considérerons que la confiance est une croyance d’un agent en l’intention d’un autre agent d’accomplir une action ou un but spécifique.

La confiance est aussi considérée dans la littérature comme un élément de *soft security* [24] qui, au niveau collectif, prévient la propagation de fausses croyances par des agents malicieux en les excluant d’interactions sociales.

Dans le cas particulier de la reconnaissance de buts, la probabilité d’un objectif donné, évaluée au regard d’une fonction de reconnaissance de but et d’un ensemble de connaissances, constitue une description d’une attente de l’agent observateur envers le comportement de l’agent observé.

De plus, par souci de concision nous ne traitons pas dans cet article de l’usage de la confiance dans le mécanisme de construction d’une coalition, mais proposons de considérer l’acquisition de la confiance des autres comme un préalable nécessaire à l’intégration d’un agent dans un collectif d’agents (pouvant prendre par exemple la forme d’une organisation) ou comme un événement déclencheur de la construction d’une telle structure. Nous nous plaçons dans le cadre de systèmes pouvant être assimilés à des jeux coopératifs (c’est-à-dire dans lesquels les agents ont

des gains ou performances accrues en cas de coopération). Ainsi, nous décrivons ici des agents qui vont poursuivre leur propre but, tout en prenant en compte l'impact social de leur comportement afin de pouvoir augmenter si possible leurs gains en gagnant la confiance d'autres agents. En plus de cet usage de l'évaluation mutuelle entre agents en vue de la coopération, nous considérons également dans le cadre de cet article l'usage opposé, c'est-à-dire consistant à employer ces mécanismes de construction de la confiance pour identifier les agents susceptibles de poursuivre des objectifs indésirables pour l'observateur. Une fois acquise la confiance dans la poursuite d'objectifs indésirables (du point de vue de l'observateur), il est possible d'appliquer des sanctions, soit unilatéralement en refusant de collaborer avec l'agent observé, soit la forme de peines décidées par un groupe d'agents pouvant réduire les gains de l'agent jugé ou l'exclure d'un collectif [17, 28]. Les sanctions associées à ces infractions peuvent être représentées explicitement sous forme d'institutions (voir par exemple le cadre conceptuel ADICO [13]) et mettre ainsi à disposition des agents une représentation compréhensible et accessible pouvant être prise en compte dans leur décision.

2.2 Reconnaissance d'objectifs

Comme mentionné en section 2.1, afin de permettre aux agents de se construire une représentation mentale de l'intention d'un autre agent, nous allons employer une fonction de reconnaissance d'objectif, que nous définirons formellement par une fonction abstraite. Être capable d'interpréter le comportement d'un autre agent fait partie intégrante de la capacité à comprendre le contexte dans lequel l'agent observateur se trouve, et plus précisément le contexte social. Diverses approches ont fait l'objet de propositions dans la littérature en Intelligence Artificielle et nous nous attachons ici, sans prétendre être exhaustifs, à montrer leur diversité et mettre en lumière les points communs.

Nous présentons en premier lieu une première catégorie, que nous nommerons *reconnaissance de plans* (PR) [20]. L'usage de plans étant classique pour décrire la manière dont les agents peuvent atteindre leurs buts au travers d'une séquence d'actions et de buts intermédiaires, cette approche regroupe les propositions d'algorithmes permettant d'évaluer les similitudes entre un comportement observé et les plans d'une bibliothèque de plans. Pour que cette approche soit efficace, la bibliothèque de plans doit être similaire sinon partagée entre l'agent observé et l'observateur. Cette approche a montré de bons résultats, même face à des agents humains ou plus généralement des cas d'observabilité partielle [11]. Au cours de l'observation, l'observateur évalue et révisé une ensemble d'*explications possibles*, dans lesquelles chaque action est associée à un plan. L'explication la plus probable, au regard du rasoir d'Ockham, est la plus simple, c'est-à-dire celle qui fait appel aux moins de plans différents possibles. L'évaluation de l'ensemble des explications souffre toutefois d'une explosion combinatoire, puisqu'il faut effectivement pour chaque action du comportement observé, envisager toutes les oc-

currences de cette action dans la bibliothèque de plans. De nombreux algorithmes proposés, tels qu'ELEXIR [10], PHATT [11], YAPPR [12], DOPLAR [15] ou encore SLIM [19] cherchent à résoudre ce problème en réduisant continuellement la taille de l'ensemble des explications possibles en procédant à l'exclusion des plus invraisemblables.

La reconnaissance de plans donne des résultats significatifs lorsque l'observateur dispose d'une bibliothèque de plans pertinente et appropriée, et si les plans la constituant sont assez détaillé et différents les uns des autres. Toutefois dans certaines applications les observateurs ne peuvent pas disposer de telles connaissances données à l'initialisation. C'est par exemple le cas d'agents employant des techniques d'apprentissage pour connaître l'effet de leurs effecteurs sur le monde perçu au travers de leur fonction de perception. C'est aussi le cas lorsque le concepteur d'un système, ou *a minima* de l'agent observateur, ne sait pas encore ce que peuvent être les comportements pertinents dans le système. Pour résoudre ce problème, certains auteurs proposent alors d'employer des techniques d'*apprentissage automatique de comportements* pour permettre à l'observateur de construire des modèles de comportement à partir d'observations. Ces modèles sont ensuite utilisés pour évaluer les comportements observés durant leur exécution. Par exemple, le domaine applicatif choisi pour présenter l'algorithme ILSA [14] montre un agent observant une personne âgée dans son domicile afin de proposer en cas de nécessité un contact avec du personnel soignant approprié. Cette approche montre tout d'abord son intérêt par sa capacité à extraire un ensemble de motifs émergeant progressivement des comportements humains observés. Sa flexibilité lui permet de composer avec les capteurs à sa disposition, les spécificités du domicile dans lequel le système est déployé et les habitudes spécifiques à la personne observée. D'autres proposent d'utiliser des techniques d'apprentissage non supervisé hors du temps d'exécution [16] pour effectuer de la recherche de motifs fréquents dans des comportements observés et révéler des descriptions abstraites de types de comportement qui seront ensuite utilisés en temps réel pour inférer les objectifs de l'agent observé durant l'exécution. Dans ce cas, le motif associé aux comportements observés une liste d'objectifs associés à une certaine probabilité extraite de la base d'apprentissage.

Les techniques présentées et regroupées dans ces deux approches –par *reconnaissance de plans* ou par *apprentissage de comportements*– proposent toutes d'informer l'observateur de l'objectif le plus probable associé à un comportement et à un ensemble de connaissances (donné à l'agent ou appris durant ou hors de l'exécution). La première catégorie est généralement préférée lorsque le concepteur dispose de connaissances sur les comportements attendus. La seconde approche, offre en revanche la possibilité de construire cette connaissance en apprenant à associer des séquences d'actions à des buts à partir d'un grand nombre d'observations. Afin de proposer une abstraction de cet ensemble de techniques, nous définissons un concept plus gé-

néral et abstrait de *fonction de reconnaissance d'objectif* ou GRF (pour *Goal Recognition Function*) telle que :

$$GRF : b_{ag_i} \times K \times G \rightarrow [0; 1]$$

où cette fonction prend en paramètres le comportement b_{ag_i} , l'ensemble de connaissances K (représenté comme une bibliothèque de plans, un ensemble de motifs ou toute autre forme de connaissances), et l'ensemble des buts G possibles associés à la probabilité $P(g_i)$ que l'agent observé ait l'intention d'atteindre le but g_i . Afin de normaliser le résultat obtenu, nous considérons par la suite la somme de toutes les probabilités associées aux objectifs comme égales à 1. En l'absence d'information (typiquement, à l'initialisation du système si l'observateur ne dispose d'aucune croyance préalable sur les intentions des autres agents, ou dans le cas où un agent est dans l'incapacité d'interpréter un comportement), cette probabilité est donc $P(g) = \frac{1}{|G|}$, $\forall g \in G$.

2.3 Gestion d'information

Puisque nous envisageons des agents capables d'employer une fonction de reconnaissance d'objectifs sur des comportements observés attribués à d'autres agents (voir Section 2.2), et employer l'information obtenue pour établir des relations de confiance (voir Section 2.1), il apparaît que le comportement d'un agent aura une influence sur ses relations avec les autres et pourrait influencer ses capacités à atteindre ses propres objectifs. L'information ainsi donnée aux autres agents au travers du comportement doit donc être prise en compte. Il est possible alors d'imaginer des agents cherchant à garder leurs intentions secrètes, pour éviter par exemple des sanctions ou un isolement, ou au contraire des agents tentant d'exposer le plus possible leurs objectifs au travers de leur comportement, afin de faciliter les coopérations et occuper un rôle central dans la société des agents autonomes artificiels.

Une stratégie de gestion de l'information ou IRS (pour *Information Regulation Strategy*) permet alors à un agent de prendre des décisions qui, en plus de lui permettre d'atteindre ses objectifs, vont permettre de minimiser ou maximiser l'information donnée. Afin de mesurer cette quantité d'information donnée à un observateur via un comportement hypothétique utilisant une fonction de reconnaissance d'objectif, l'agent calcule l'entropie de Shannon [27] de l'ensemble des probabilités que cette fonction attribue aux objectifs possibles. L'entropie $H_{b,K}$ correspondant au comportement b au regard d'un ensemble de connaissances K est obtenue grâce à la formule suivante :

$$H_{b,K} = - \sum_{g \in G} P(g | b, K) \cdot \log(P(g | b, K))$$

Deux stratégies de gestion de l'information sont alors envisageables : l'agent peut adopter une stratégie de *transparence*, c'est-à-dire minimiser l'entropie sur la distribution des probabilités afin de rendre son objectif réel le plus évident possible, ou au contraire adopter une stratégie de *dissimulation*, en maximisant l'entropie afin de maintenir

l'observateur dans un état de confusion le plus longtemps possible. Une stratégie de gestion de l'information est donc mise en œuvre au niveau du processus de prise de décision d'un agent, en planifiant ses actions non seulement pour atteindre les objectifs, mais également pour atteindre et maintenir une certaine certitude ou incertitude chez un observateur. Appliquer une telle stratégie peut nécessiter d'effectuer des actions qui seraient, en l'absence d'un tel critère de choix, considérées comme sous-optimales (plus longues, nécessitant plus de ressources, etc.). Nous parlons alors de coûts de mise en œuvre de ces stratégies.

Notons à ce stade que la stratégie de dissimulation est à distinguer de la définition la plus intuitive de mensonge ou de tromperie, qui consisterait à maximiser la probabilité associée à un autre but que celui réellement poursuivi par l'agent. La dissimulation ne cherche pas à créer de fausse croyance chez l'observateur, en l'amenant à croire que le comportement de l'agent vise à réaliser un autre but, mais seulement à maintenir une situation ambiguë et une absence d'information [5].

Après avoir montré comment des agents peuvent bâtir des relations de confiance à partir de diverses méthodes de reconnaissance d'objectifs, et introduit le concept de gestion d'information, il est nécessaire d'expliquer comment un agent peut choisir la stratégie appropriée compte tenu du contexte social et de ses propres objectifs. Nous allons maintenant nous attacher à montrer dans la suite de cet article en Section 3 comment employer ces stratégies de gestion de l'information en fonction du contexte social perçu par l'agent afin de maximiser ses espérances de gain (et minimiser les risques de sanctions).

3 Sélectionner une stratégie de gestion de l'information dans un système de confiance

Comme expliqué précédemment, nous proposons cette contribution comme un moyen de sélectionner la stratégie de gestion de l'information la plus appropriée compte tenu du contexte. Par contexte, nous entendons principalement les états mentaux produits par l'évaluation du comportement des autres (images et confiances, donc).

3.1 Évaluation de l'utilité des stratégies

Tout d'abord, le processus de sélection va évaluer l'utilité espérée de l'une ou l'autre stratégie dans le contexte présent selon les formules suivantes :

$$\begin{aligned} EU_T &= P_{C|T} \cdot Benefit_C - Cost_T \\ EU_O &= P_{C|O} \cdot Benefit_C - Cost_O \end{aligned}$$

Où l'utilité espérée d'une stratégie de transparence, EU_T (pour *Expected Utility of transparency*), dépend de la probabilité $P_{C|T}$ d'être inclus dans un collectif d'agents C en suivant cette stratégie, les bénéfices $Benefit_C$ d'une telle inclusion et le coût $Cost_T$ associés à la révélation de ses objectifs (ou perte des bénéfices liés à la confidentialité) et

incluant les coûts de mise en œuvre de cette stratégie, ainsi que d'éventuelles sanctions.

De manière analogue, l'utilité espérée d'une stratégie de dissimulation, EU_O (pour *Expected Utility of obfuscation*), dépend de la probabilité $P_{C|O}$ d'être inclus dans ce même collectif C tout en suivant cette stratégie et le coût $Cost_O$ de la mise en œuvre de la stratégie.

Trois cas de figures peuvent alors se présenter :

- si $EU_T > EU_O$, l'agent a alors intérêt à adopter une IRS de transparence
- si $EU_T < EU_O$, l'agent a alors intérêt à adopter une IRS de dissimulation
- si $EU_T = EU_O$, l'agent sait alors explicitement que les deux stratégies sont d'intérêt équivalent (il aura alors besoin d'une règle de bris d'égalité pour sélectionner l'une ou l'autre).

Les définitions précises des probabilités, coûts et bénéfices dépendent du contexte applicatif et ces éléments du modèle sont présentés ici de manière volontairement abstraite. Un exemple d'implémentation triviale mais concrète de ces fonctions est proposé en Section 4.1. Nous considérons à ce stade par simplicité que l'agent connaît leur valeur à travers d'une croyance. Il est par exemple envisageable de laisser l'agent réévaluer la probabilité $P_{C|T}$ d'être admis au sein d'un collectif en observant la proportion d'agents effectivement admis parmi ceux identifiés comme ayant eux aussi adopté cette IRS.

3.2 Sélection de la stratégie de régulation

Après avoir défini en Section 3.1 les modalités d'évaluation de l'utilité des stratégies de dissimulation et de transparence, nous illustrons ici comment l'agent sélectionne la stratégie la plus adaptée en fonction de sa perception du contexte, et plus particulièrement du contexte social.

Nous considérons ici les stratégies de gestion d'information évoquées en Section 2.3. Nous ajoutons à cela une stratégie appelée *indifférent*, consistant à ne pas tenir compte de la quantité d'information transmise à l'observateur lors de la prise de décision. Cette troisième stratégie sera employée lorsque l'agent considère que la gestion d'information n'est pas un critère pertinent dans ses interactions sociales pour influencer la décision des autres dans la formation de collectifs d'agents. Nous introduisons cette croyance $BsocImpact$ signifiant, lorsqu'elle est présente dans l'ensemble des croyances B de l'agent, que l'agent est convaincu que l'information transmise à un observateur *via* son comportement aura un impact social (c.-à-d. que la connaissance issue de l'observation de son comportement influence la décision des autres agents de l'inclure dans des coalitions ou de l'en exclure). La production et révision de cette croyance ne sera pas traitée en détail dans cet article mais par exemple, dans un système normatif, il est envisageable que l'agent vérifie au préalable si la poursuite de son objectif est acceptable ou non au regard de normes et d'institutions [8]. De même, il est envisageable que l'agent, s'il venait à être doté d'une fonction de reconnaissance de stratégie (voir Section 5), puisse évaluer si l'emploi d'une stratégie particulière est corrélée à une meilleure inclusion

ou exclusion.

Lors de la mise à jour de sa base de croyances, l'agent évalue la stratégie la plus appropriée en fonction de ses croyances et utilités espérées pour les différentes stratégies.

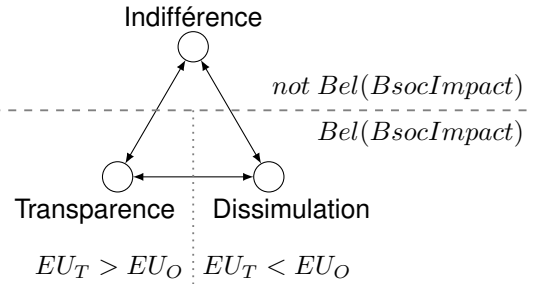


FIGURE 1 – Transitions entre stratégies

La figure 1, illustre la manière dont un agent peut évoluer d'une IRS à l'autre en fonction des utilités espérées et de ses croyances. Premièrement, seul un agent ayant la croyance $BsocImpact$ a intérêt à adopter une autre stratégie que la stratégie d'*indifférence*, les notions d'utilité espérée n'ayant aucune pertinence dans un tel contexte. Ensuite, la comparaison des utilités permet à un agent de basculer d'une stratégie de *transparence* à une stratégie de *dissimulation* et réciproquement. Le schéma ne fait volontairement pas mention de la règle de bris d'égalité discutée en section 3.1

4 Expérimentations

Nous avons fait le choix dans cet article de fournir un modèle qui propose qu'une définition volontairement abstraite de certains éléments (fonctions d'utilité, règle de bris d'égalité, etc.) en laissant au concepteur d'un réel système la tâche de les définir. Nous justifions ce choix par la nécessité de s'adapter à la grande diversité des domaines applicatifs et mécanismes de reconnaissance de but présentés en Section 2. Cette section présente un exemple arbitraire mais simple de mise en œuvre du modèle afin d'illustrer les effets de l'emploi de la contribution proposée dans cet article et explorer les limites de notre proposition. Nous avons choisi de présenter ici les résultats obtenus à la suite d'une démarche d'expérimentation. Pour cela, nous présenterons brièvement l'implémentation de ce modèle à l'aide du cadriciel JaCaMo [1] en Section 4.1. Ensuite nous illustrons les mécanismes présentés dans cet article au travers des résultats d'exécution d'un scénario en Section 4.2.

4.1 Mise en œuvre

Description du système

Nous avons choisi d'implémenter cette preuve de concept à l'aide du cadriciel multi-agent JaCaMo [1], connu dans la communauté pour offrir une architecture inspirée du modèle BDI [23] dont les croyances, plans et raisonnements peuvent être décrits à l'aide du langage AgentSpeak [3] dont la syntaxe est proche du paradigme de programmation logique. Jason, l'implémentation d'AgentSpeak proposée dans JaCaMo, propose de définir les plans

de l'agent en associant à un évènement déclencheur (ou *triggering event*), un contexte d'activation (une expression logique évaluée au moment de la sélection de plans), et *corps* formé d'une séquence d'actions internes (e.g. ajout/suppression de croyances, déclenchement d'évènements, ajout/suppression de buts) et externes (e.g. communication, interactions avec l'environnement). La description de l'environnement se fait à l'aide de la bibliothèque CArAgO [25] en langage Java et consiste principalement en la description d'*artefacts*, proposant aux agents des actions et pouvant informer ces agents via un système d'évènements. Dans l'expérimentation présentée ici, les agents sont tous connectés à un même artefact, proposant douze actions différentes $\{a_1, \dots, a_{12}\}$ permettant chacune de progresser vers un ou plusieurs objectifs. Tous les agents sont informés par l'artefact de l'exécution d'une action et de l'identité de son auteur lorsqu'elle est exécutée. Afin de représenter d'éventuelles imperfections dans la fonction de perception ou la fonction de reconnaissance d'objectifs, l'artefact possède un attribut $pfail \in [0; 1]$ modélisant la probabilité d'une erreur de perception et/ou d'interprétation par la GRF (voir Section 2.2), et aboutissant à un bruit plus ou moins significatif dans la construction des images, permettant de modéliser un environnement partiellement observable. Une valeur non-nulle de $pfail$ permet par exemple de modéliser la difficulté des agents à distinguer correctement des actions effectuées dans l'environnement. Cela peut venir par exemple d'erreurs de capteur, de problèmes de perception locale ou encore d'un manque d'occurrences de certains comportements dans une base d'apprentissage.

Les agents vont, à partir des images construites au cours de leurs observations du comportement des autres agents, accorder et retirer leurs confiances dans les objectifs des autres. Nous avons ici adopté un modèle simple dans lequel les probabilités fournies par la GRF sont utilisées comme images et à partir desquels la confiance est établie lors de dépassement d'une valeur seuil fixée à $P(G) = 0.6$. Notons que, la somme des probabilités étant normalisée à 1, il est impossible d'avoir ici confiance en l'intention d'un agent d'accomplir plusieurs objectifs simultanément.

Implantation des IRS

Les IRS, dans le cadre de cette implémentation, ont été introduites dans ce modèle sous forme de prédicats utilisés dans le contexte de plans. Ainsi, ces IRS sont employées lors de la sélection de plan afin de choisir celui qui (exception faite de l'*indifférence*) minimise ou maximise l'entropie dans la distribution des probabilités des objectifs.

Les objectifs des agents, l'IRS courante, les comportements observés des autres agents et les images et confiances construites à partir de ces observations sont représentés sous forme d'états mentaux dans la base de croyance. Ces états mentaux sont révisés au fur et à mesure de la perception successive d'actions exécutées dans l'environnement.

Pour observer le comportement du système, nous observerons les états mentaux d'un agent du système, et plus parti-

culièrement l'évolution des probabilités associées aux objectifs possibles des agents, l'entropie de cette distribution de probabilités (représentant donc une mesure de la quantité d'information détenue par l'observateur choisi), et les utilités associées aux différentes stratégies.

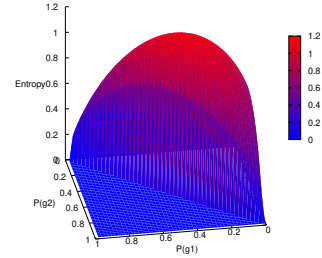


FIGURE 2 – Représentation de l'entropie dans un espace tridimensionnel décrit par les probabilités associées à trois buts différents.

La Figure 2, propose une représentation de l'entropie en fonction de la probabilité des buts. Ici, l'espace de toutes les distributions possibles des probabilités est projeté sur un plan, et la dimension verticale représente l'entropie $H_{b,L}$ correspondant à ces distributions. L'entropie est maximale pour $P(g_1) = P(g_2) = P(g_3) = \frac{1}{3}$, c'est-à-dire lorsque l'information détenue par l'observateur est minimale. À l'inverse, lorsque $\exists g \in G : P(g) = 1$, l'information est maximale et l'entropie associée $H_{b,L} = 0$. Comme on peut ainsi le voir sur cette figure, adopter une stratégie de *dissimulation*, telle que décrite formellement en Section 2.3 revient à rester autant que possible proche du sommet de cette fonction, où la confusion est maximale pour l'observateur. La stratégie de *transparence* en revanche consiste à adopter le comportement qui fera migrer cette entropie le plus vite possible vers le "pôle" correspondant à la probabilité maximale pour le but réel de l'agent, et en rester le plus proche possible. à ces pôles, l'entropie atteint une valeur nulle.

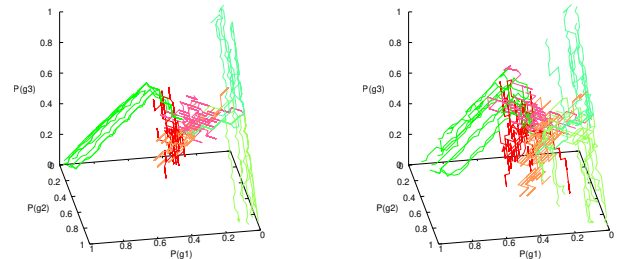


FIGURE 3 – Évolutions des probabilités des objectifs pour $pfail=0.1$ (à gauche) et $pfail=0.2$ (à droite)

La Figure 3 illustre du point de vue de l'un des agents du système, l'évolution des images de 60 agents dans un espace où chaque dimension correspond à la probabilité es-

timée de l'un des buts. Les probabilités associées à leurs objectifs sont initialisées à $P(g)_{t_0} = \frac{1}{|G|}$, point correspondant en Figure 2 au sommet de la courbe, où l'entropie est maximale.

En couleurs chaudes, des agents ayant adopté une IRS de dissimulation tendent à rester le plus loin possible des pôles, c'est-à-dire là où l'entropie sur les probabilités est maximale. En revanche, les agents ayant adopté une IRS de *transparence* migrent vers les trois pôles de l'intersection formée entre l'espace et le plan défini par l'équation $P(g1)+P(g2)+P(g3) = 1$, comme défini en Section 2.2. Plus la probabilité *pfail* d'erreur de perception et d'interprétation est importante, plus les agents ayant un même objectif et une même IRS sont dispersés sous l'effet des erreurs de perception. Cette dispersion peut aller jusqu'à conduire un observateur à accorder par erreur sa confiance à un agent.

Selection d'une IRS

Le processus d'évaluation des IRS est appelé à chaque perception de l'agent, vérifiant l'existence ou l'absence de la croyance *BsocImpact* et révisant, s'il y a lieu, les utilités espérées pour les stratégies de *transparence* et de *dissimulation*. Nous proposons dans le cadre de cette implémentation, d'évaluer les utilités espérées à l'aide des valeurs suivantes :

$$Benefit_C = \sum_{ag \in Ag} benefit(ag)$$

avec

$$benefit(ag) = \begin{cases} \frac{1}{|Ag|} ssi \ trust(ag, g) \wedge myGoal(g) \\ 0 \ sinon \end{cases}$$

ainsi que

$$Cost_T = \sum_{ag \in Ag} cost_T(ag)$$

avec

$$cost_T(ag) = \begin{cases} \frac{1}{|Ag|} ssi \ trust(ag, g) \wedge not \ myGoal(g) \\ \frac{1}{2 \cdot |Ag|} ssi \ \nexists g \in G \ s.t. \ trust(ag, g) \\ 0 \ sinon \end{cases}$$

et

$$Cost_O = 0$$

$$P_{C|T} = 0.9$$

$$P_{C|O} = 0.1$$

Avec de telles définitions des fonctions de coûts et bénéfices, les agents qui n'adoptent pas une IRS d'indifférence seront incités à dissimuler leurs objectifs tant qu'ils n'ont pas confiance en une proportion suffisante d'agents pour poursuivre des objectifs similaires au siens.

4.2 Résultats

Nous proposons d'étudier le comportement du système au travers d'un premier scénario. Dans ce scénario, deux agents *ir1* et *ir2* sont initialisés avec pour objectifs respectifs *g1* et *g2*, et la croyance *BsocImpact* (voir Section 3.2). Deux autres agents *ir3* et *ir4*, sont initialisés quand à eux avec un même objectif *g1*, mais sans la croyance *BsocImpact*. Les agents sont initialisés sans *a priori* sur les buts des autres agents du système. Le même scénario est joué à de multiples reprises afin d'observer les propriétés du système, peu importe les situations particulières créées par les aléas venant de l'absence d'ordonnancement dans la plateforme et les erreurs de perceptions des agents.

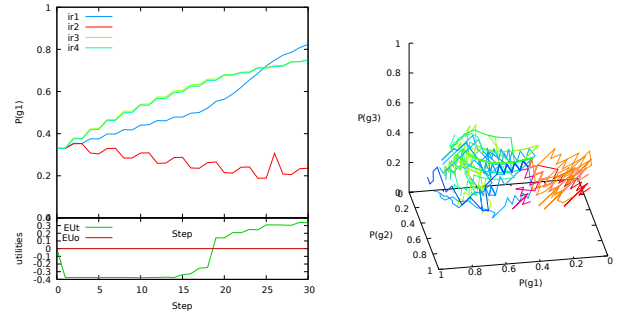


FIGURE 4 – Observation de l'évolution des images des agents associés au but *g1* et des utilités espérées (figure de gauche), et évolution des probabilités associées aux divers buts possibles (figure de droite), du point de vue de l'agent *ir1*. Résultats expérimentaux agrégés sur 100 runs, avec *pfail*=0.05.

La Figure 4 montre le point de vue de l'agent *ir1* et l'évolution de sa vision des autres agents au cours de la simulation. Comme l'agent ne dispose pas d'informations au départ sur les objectifs des autres agents, l'utilité espérée d'une IRS de transparence tombe dès les premières observations à un niveau inférieur à celle de l'IRS de dissimulation.

Les agents *ir3* et *ir4*, qui ne disposent pas de la croyance *BsocImpact*, adoptent un comportement qui va accroître de manière régulière la probabilité associée à leur objectif, tandis que pour *ir1* et *ir2*, l'usage d'une IRS de dissimulation leur permet de ne donner que peu d'informations aux observateurs (ce qui se traduit par un accroissement moindre de la probabilité $P(g1)$ durant les premières actions de la simulation).

Lorsque les images des agents *ir3* et *ir4* sont assez significatives pour permettre à l'agent *ir1* d'accorder sa confiance, l'utilité de l'IRS de transparence remonte brutalement en raison de l'intérêt perçu par *ir1* de dévoiler son objectif à des observateurs qui semblent majoritairement le partager. L'agent change alors sa stratégie (le changement est explicitement mentionné dans les logs). L'image que l'agent *ir1* construit à partir de l'observation de son propre comportement entame alors une brusque ascension.

Pour l'agent *ir2* en revanche, les agents *ir3* et *ir4*, puis *ir1* étant identifiés comme poursuivant un autre objectif que le sien, l'IRS de dissimulation reste celle ayant l'uti-

lité la plus élevée. La figure de droite montre comment, au cours des multiples itérations de l'exécution de ce scénario, l'agent ir_2 –représenté en couleurs chaudes– tend à rester au centre de l'espace des probabilités, où l'entropie est maximale (comme illustré en Figure 2, Section 4.1).

Ce scénario illustre comment la sélection d'IRS, via les définitions d'utilités exposées ici, permet de décrire des agents prudents et opportunistes, c'est-à-dire attendant d'avoir obtenu suffisamment d'information sur le comportement des autres agents du système pour décider ou non de dévoiler leurs propres objectifs au travers de leur comportement.

5 Conclusion

Au travers de cet article, nous avons tout d'abord rappelé les enjeux de la prise en compte de l'information transmise aux observateurs dans des systèmes où la confiance des agents envers les autres peut impacter leur performance. Nous avons montré comment la littérature propose de doter les agents de fonctions de reconnaissance d'objectifs à partir de comportements observés et inférer ainsi l'objectif le plus vraisemblable d'un agent au regard d'un ensemble de connaissances. Nous avons ensuite montré comment la gestion d'information permet à un agent de prendre connaissance de la quantité d'information transmise à un observateur via son comportement, et comment anticiper cela dans sa prise de décision à l'aide d'une stratégie de gestion de l'information. Cet article a ensuite présenté une proposition de mécanisme de sélection de la stratégie de gestion de l'information en fonction du contexte, et plus particulièrement des relations de confiance établies avec les autres agents à partir de ses propres observations. Ce mécanisme de sélection repose sur une évaluation des utilités des stratégies à disposition en fonction des bénéfices espérés en dévoilant ses objectifs, comparés aux coûts associés à d'éventuelles sanctions en réaction à une telle révélation. Enfin, une preuve de concept illustre la mise en œuvre de ce modèle à travers un système implémenté et testé au travers d'un scénario dans lequel les agents vont employer ce processus pour changer de stratégie en fonction de l'évolution de la situation.

Un tel modèle présente tout son intérêt dans les domaines applicatifs où les agents peuvent poursuivre plusieurs buts différents et ont des raisons valables de ne pas désirer voir certains autres agents accomplir certains buts (par exemple pour des motivations stratégiques ou éthiques) et où les agents peuvent avoir des interactions sociales conditionnées par une certaine confiance (formation de coalitions, construction d'organisation, partage de ressources, coordination, partage d'information, etc.). Par exemple, un tel modèle pourrait être pertinent dans le cadre de gestion éthique d'actifs financiers [6], où des agents dotés de théories morales différentes permettant de définir ce qu'est un investissement éthique cherchent à déterminer les partenaires appropriés pour des échanges de gré-à-gré.

Les résultats expérimentaux montrent comment, dans un cadre abstrait et simple, les agents peuvent adapter leur

stratégie à l'évolution de leur image des autres agents, et ainsi faire preuve d'une forme de prudence dans la révélation de leurs objectifs. Toutefois, nous aimerions voir ces résultats confirmés dans le cadre d'un déploiement dans un domaine applicatifs réaliste, et pourquoi pas avec des agents humains (dans un cadre vidéoludique par exemple).

Ce modèle repose également sur de nombreuses hypothèses : l'hypothèse de l'efficacité des perceptions et de la fonction de reconnaissance d'objectif par exemple. Il serait intéressant à ce titre d'explorer les effets d'une hétérogénéité de points de vue ou de fonctions implémentées, pour modéliser par exemple le cas d'agents dotés de capteurs, d'architectures ou de raisonnements différents.

De plus, nous avons proposé ici un modèle de sélection de la stratégie de gestion de l'information en nous reposant sur un système de confiance, et il serait intéressant de l'étendre à la prise en compte de systèmes de réputation [26] en laissant les agents membres d'un collectif transmettre leurs états mentaux. Une telle extension serait d'autant plus intéressante pour les contextes applicatifs dans lesquels les agents ne disposent que d'une perception locale et donc avoir des points de vues divergeant sur un même agent en raison de leurs expériences individuelles.

Enfin, de manière analogue à la *fonction de reconnaissance de buts* utilisée par les agents, nous aimerions proposer une *fonction de reconnaissance de stratégies*, permettant à un agent de reconnaître avec une certaine probabilité la stratégie de gestion de l'information employée par un autre agent. Une telle fonctionnalité ouvrirait la voie à la construction de confiance sur la gestion de l'information et permettrait par exemple de faire émerger des groupes d'agents qui, en plus de se considérer mutuellement comme cherchant à atteindre des buts acceptables, font preuve d'une transparence qui serait également appréciée. De tels modèles pourraient également présenter un intérêt pour accorder des privilèges dans des systèmes ayant de forts enjeux de sécurité en réservant l'accès à certaines ressources aux agents dont l'objectif est identifié et accepté, et semblant adopter une stratégie de transparence. À l'inverse il est possible d'imaginer aussi que l'emploi efficace d'une stratégie de dissimulation soit valorisée dans certains groupes, par exemple lorsque la réussite des objectifs du groupe est conditionnée par l'ignorance des agents exclus du collectif (par exemple dans le cas de jeux de stratégie ou pour modéliser des problèmes de *privacy*).

Informations complémentaires et reproductibilité des résultats

L'implémentation de la preuve de concept, des informations complémentaires sur les mécanismes implémentés et les données employées pour produire les figures sont mis à disposition sur un dépôt en ligne. Il est également possible de créer de nouveaux scénarios en modifiant les agents et leurs croyances initiales. L'ensemble peut être téléchargé en clonant le dépôt git à l'adresse suivante : <https://gitlab.com/NCointe/poc-jfisma20>

Remerciements

Ce travail a été réalisé grâce au support financier du Conseil Européen de la Recherche (ERC) au sein du programme recherche et innovation Horizon 2020 de l'Union Européenne via le projet BEHAVE (numéro d'agrément 724431).

Références

- [1] Olivier Boissier, Rafael H Bordini, Jomi F Hübner, Alessandro Ricci, and Andrea Santi. Multi-agent oriented programming with jacamo. *Science of Computer Programming*, 78(6) :747–761, 2013.
- [2] Grégory Bonnet. A protocol based on a game-theoretic dilemma to prevent malicious coalitions in reputation systems. In *ECAI 2012 - 20th European Conference on Artificial Intelligence. Including Prestigious Applications of Artificial Intelligence (PAIS-2012) System Demonstrations Track, Montpellier, France, August 27-31, 2012*, pages 187–192, 2012.
- [3] Rafael H Bordini, Jomi Fred Hübner, and Michael Wooldridge. *Programming multi-agent systems in AgentSpeak using Jason*, volume 8. John Wiley & Sons, 2007.
- [4] Zack Butler, Robert Fitch, and Daniela Rus. Distributed control for unit-compressible robots : goal-recognition, locomotion, and splitting. *IEEE/ASME transactions on mechatronics*, 7(4) :418–430, 2002.
- [5] CG Chorus. How to keep your av on the right track ? an obfuscation-based model of decision-making by autonomous agents. *hEART 2018*, 2018.
- [6] Nicolas Cointe, Grégory Bonnet, and Olivier Boissier. Ethics-based cooperation in multi-agent systems. In *14th Social Simulation Conference (SSC18)*, August 2018.
- [7] Nicolas Cointe, Amineh Ghorbani, and Caspar Chorus. Management of obfuscation-based decision making in a coalition. In *15th Social Simulation Conference (SSC19)*, September 2019.
- [8] Virginia Dignum and Julian Padget. *Multiagent organizations*, chapter 2, pages 51–98. MIT press, second edition, 2013.
- [9] Diego Gambetta et al. Can we trust trust. *Trust : Making and breaking cooperative relations*, 13 :213–237, 2000.
- [10] Christopher W Geib. Delaying commitment in plan recognition using combinatory categorial grammars. In *IJCAI*, pages 1702–1707, 2009.
- [11] Christopher W Geib and Robert P Goldman. Partial observability and probabilistic plan/goal recognition. In *Proceedings of the International workshop on modeling other agents from observations (MOO-05)*, volume 8, pages 1–6, 2005.
- [12] Christopher W Geib, John Maraist, and Robert P Goldman. A new probabilistic plan recognition algorithm based on string rewriting. In *ICAPS*, pages 91–98, 2008.
- [13] Amineh Ghorbani and Giangiacomo Bravo. Managing the commons : a simple model of the emergence of institutions through collective action. *International Journal of the Commons*, 10(1), 2016.
- [14] Valerie Guralnik and Karen Zita Haigh. Learning models of human behaviour with sequential patterns. In *Proceedings of the AAAI-02 workshop "Automation as Caregiver"*, pages 24–30, 2002.
- [15] Froduald Kabanza, Julien Filion, Abder Rezak Benaskeur, and Hengameh Irandoust. Controlling the hypothesis space in probabilistic plan recognition. In *IJCAI*, pages 2306–2312, 2013.
- [16] Andreas D Lattner, Andrea Miene, Ubbo Visser, and Otthein Herzog. Sequential pattern mining for situation and behavior prediction in simulated robotic soccer. In *Robot Soccer World Cup*, pages 118–129. Springer, 2005.
- [17] Samhar Mahmoud, Simon Miles, and Michael Luck. Cooperation emergence under resource-constrained peer punishment. In *AAMAS*, pages 900–908, 2016.
- [18] D Harrison McKnight and Norman L Chervany. What trust means in e-commerce customer relationships : An interdisciplinary conceptual typology. *International journal of electronic commerce*, 6(2) :35–59, 2001.
- [19] Reuth Mirsky et al. Slim : Semi-lazy inference mechanism for plan recognition. *arXiv preprint arXiv :1703.00838*, 2017.
- [20] Reuth Mirsky, Roni Stern, Kobi Gal, and Meir Kalech. Sequential plan recognition : An iterative approach to disambiguating between hypotheses. *Artificial Intelligence*, 260 :51–73, 2018.
- [21] G. Muller, L. Vercoouter, and O. Boissier. Towards a general definition of trust and its application to openness in MAS. In *6th Workshop on Deception, Fraud and Trust in Agent Societies*, pages 49–56, 2003.
- [22] Jérémy Patix, Abdel-Allah Mouaddib, Simon Le Gloannec, Dafni Stampouli, and Marc Contat. Discrete relative states to learn and recognize goals-based behaviors of groups. In *Proceedings of the 2013 international conference on Autonomous agents and multi-agent systems*, pages 933–940. International Foundation for Autonomous Agents and Multiagent Systems, 2013.
- [23] Anand S Rao, Michael P Georgeff, et al. BDI agents : from theory to practice. In *ICMAS*, volume 95, pages 312–319, 1995.
- [24] Lars Rasmusson and Sverker Jansson. Simulated social control for secure internet commerce (position paper). In *Proceedings, New Security Paradigms Workshop, Lake Arrowhead*, 1996.

- [25] Alessandro Ricci, Michele Piunti, Mirko Viroli, and Andrea Omicini. Environment programming in CArtAgO. In *Multi-agent programming*, pages 259–288. Springer, 2009.
- [26] Jordi Sabater-Mir and Laurent Vercouter. Trust and reputation in multiagent systems. *Multiagent systems*, page 381, 2013.
- [27] CE Shannon and W Weaver. *The mathematical theory of communications*. University of Illinois press. Urbana. 1963.
- [28] Subhasis Thakur, Guido Governatori, and Abdul Sattar. On modeling punishment in multi-agent systems. In *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems-Volume 3*, pages 1375–1376. Citeseer, 2012.