

Révision des croyances dans un environnement multiagent : Une approche basée sur la crédibilité et les arguments

Imed Jarras, Brahim Chaib-draa

Université Laval, Département d'informatique,
Québec , QC, G1K 7P4, Canada
jarras@ift.ulaval.ca, chaib@ift.ulaval.ca

Résumé Peu de recherches se sont penchées sur la problématique de la révision des croyances dans un cadre multiagent. En tout cas, à notre connaissance, aucune ne s'est penchée sur la révision des croyances tenant compte de la crédibilité des informateurs, tout en gardant trace des arguments en faveur de la révision, une fois celle-ci effectuée. C'est ce problème qui nous a motivé et pour lequel, nous proposons ici, une approche formelle basée sur la logique étiquetée.

1 Introduction

AGM (du nom de ses auteurs Alchourrón, Gärdenfors et Makinson) est, sans aucun doute, l'une des théories de changement des croyances les mieux élaborées et les plus influentes. En plus de donner des assises théoriques solides au changement de croyances, AGM projette la lumière sur plusieurs aspects de la dynamique des croyances en général. Cette théorie a été développée par Alchourrón, Gärdenfors et Makinson [AGM85], [AM85], [Gär88]. Cependant, la théorie AGM ne tient pas compte des interactions entre agents et leurs effets possibles sur le changement des croyances. Plusieurs autres aspects qui nous semblent importants, ne sont pas considérés par AGM ou le sont d'une manière inappropriée pour un contexte multiagent. Ainsi, par exemple, l'expansion ou l'acceptation de nouvelles informations est considérée dans AGM comme une opération simple qui se résume à l'ajout de ces informations à l'ensemble des croyances. Dans un environnement multiagent, un agent doit-il nécessairement adopter cette manière de traiter les informations qu'il reçoit ? Ne pourrait-il pas juger le rejet de ces nouvelles informations comme préférable à leur acceptation ? La même question se pose pour la contraction et la révision : ne serait-il pas permis pour un agent autonome de persister dans son état de croyance s'il considère que persister est plus raisonnable que d'effectuer un changement de croyances ? Plusieurs autres questions se posent concernant la révision des croyances dans

⁰ Ce travail est supporté par le Conseil de Recherche en Sciences et en Génie du Canada (CRSNG), et par le Conseil de Recherche en Sciences Humaines du Canada (CRSH).

⁰ *Toute Correspondance doit être adressée à* : B. Chaib-draa, Département d'Informatique, Université Laval, Québec, QC, G1K 7P4 Canada, chaib@ift.ulaval.ca.

un contexte multiagent. AGM ne donne pas de réponses à ce type de questions. Nous considérons qu'il est nécessaire d'intégrer de nouveaux aspects du changement des croyances pour répondre à ces questions spécifiques à un contexte multiagent. Ce sont ces différents points qui nous ont motivés pour entamer cette étude que nous présentons succinctement dans ce papier.

2 Théorie fondationnelle et Théorie cohérentielle

Dans les approches basées sur la cohérence, lorsqu'un agent doit changer de croyances, il doit maintenir la consistance de ses états mentaux, tout en assurant un minimum de changement. Dans cette optique, l'agent conserve une croyance aussi longtemps qu'elle est cohérente avec ses autres croyances, et il doit conserver le plus de croyances possibles. Une des propriétés des approches cohérentielles réside dans le fait que les croyances cohérentes se supportent mutuellement.

Dans l'approche fondationnelle au contraire, un agent dérive ses croyances sur la base des justifications (ou de *raisons*). Ainsi, un agent ne retient une croyance dans son état mental que si, et seulement si, cet agent possède une raison, ou une justification satisfaisante pour la conserver. Dans ce cas, changer de croyance revient donc, dans cette approche, à supprimer celles qui ne sont plus soutenues par des justifications, et à ajouter des croyances qui viennent de l'être. Dans le même ordre d'idée, si un agent trouve que la justification ou la raison d'une croyance p est plus "forte" que la justification pour $\neg p$, l'agent aura tendance à réviser ses croyances et à remplacer p par $\neg p$.

Une logique étiquetée comme la LDS (Labelled Deductive Systems) de Gabbay [Gab96] est assez riche pour rendre compte des deux approches, mais on conviendra qu'elle s'y prête très bien pour une approche fondationnelle, dans la mesure où l'unité déclarative peut nous servir pour représenter le ou les justificatifs en faveur d'une formule donnée. Ainsi, la formulation générale adoptée par Gabbay du genre $a : p$ (pour dire que l'argument "a" supporte la formule p) pourrait être généralisée à :

$$\{a_1, a_2, \dots, a_n\} : p$$

de manière à signifier que la vérité de p est supportée par les justifications $\{a_1, a_2, \dots, a_n\}$. Par exemple, les éléments suivants $\{p, p \rightarrow o, o \rightarrow v\}$ supportent v . Ainsi, la logique étiquetée LDS nous permet de représenter :

$$\{p, p \rightarrow o, o \rightarrow v\} : v$$

Une telle représentation est en fait un *argument* pour lequel $\{p, p \rightarrow o, o \rightarrow v\}$ est appelé *support* de l'argument et v sa conclusion.

Nous reviendrons plus tard sur ces aspects d'arguments, auparavant il convient de donner quelques indications sur le changement de croyances dans un système multiagent (SMA).

3 Changement de croyances dans un SMA

Le changement de croyances tel qu'il est défini dans AGM concerne un agent isolé. Cet isolement ne signifie pas l'absence de contact avec le monde car les entrées (nouvelles informations) ne sont, en fait, que les résultats de ce contact. L'isolement dont on parle se traduit uniquement par le fait qu'aucun facteur externe, à part les entrées épistémiques, n'est pris en considération dans la révision des croyances : la structure logique de l'état de croyance d'un agent ne dépend pas de ses contacts avec le monde. Cela provient, en quelque sorte, du fait qu'aucune typologie n'est proposée par AGM pour rendre compte des différents types d'entrées. Dans AGM, ces dernières ne sont définies qu'en vertu des effets qu'elles engendrent sur l'état de croyance. Ceci ne permet pas à l'agent concerné par le changement des croyances de distinguer, par exemple les informations obtenues par ses propres observations des informations reçues des autres et concernant leurs observations.

Un autre point qui nous paraît important, a été négligé par AGM. Ce point concerne le fait qu'AGM définit les opérations de changement des croyances qui permettent à un agent de réviser ses croyances d'une manière rationnelle mais ne donne aucune réponse concrète à la manière de choisir le type d'opération de changement appropriée à la situation de l'agent. Prenons l'exemple suivant : supposons qu'un agent reçoit une nouvelle information qui est inconsistante avec son état de croyance. Dans ce cas, l'agent peut faire une révision ou une contraction. Mais laquelle des deux doit-il faire ? AGM ne donne aucun moyen pour choisir entre les deux.

Considérons maintenant le changement des croyances dans un SMA. Dans ce nouveau contexte, nous considérons que les agents du système ont les propriétés suivantes :

- *La sociabilité* : chaque agent interagit avec les autres agents du système. Ces interactions peuvent avoir plusieurs formes. L'une des plus importantes est la communication. Chaque agent peut envoyer de l'information aux autres agents du système comme il peut en recevoir d'eux. Ces informations peuvent affecter les croyances des agents qui les reçoivent.
- *L'autonomie* : chaque agent a ses propres croyances. Celles-ci définissent la représentation que se fait l'agent de son environnement. En plus d'être privée, cette représentation peut être tout à fait différente de la réalité de l'environnement et des représentations des autres agents. En plus, elle ne cerne pas tous les aspects de l'environnement. En d'autres termes, elle est incomplète et imparfaite, et par conséquent sujette à révision. Cependant, chaque agent a le contrôle total sur ses croyances. Aucun autre agent ne peut intervenir dans le déroulement du processus de révision ni entraver son accomplissement.

Les idées de base sur lesquelles on va aborder le changement de croyances dans un SMA sont :

- *La consistance* : chaque agent doit avoir un état de croyance consistant. La consistance du système multi-agent est considérée à un niveau local. Cette

hypothèse s'explique par le fait que chacun des agents doit maintenir la consistance de ses croyances lorsqu'il effectue une révision.

- *La minimalité du changement* : en révisant ses croyances, chaque agent fait le minimum de changements possibles. La minimalité du changement est considérée d'un point de vue qualitatif et elle est relative à chaque agent.
- *L'acceptation* : nous considérons, comme dans AGM, que les croyances sont toutes certaines mais changeables. Cependant, la vulnérabilité au changement diffère d'une croyance à une autre.
- *La persistance* : un agent n'est pas obligé d'intégrer impérativement les nouvelles informations. Il peut persister sur ses attitudes de croyance. Ceci est valable, spécialement, pour les informations qu'il reçoit des autres agents. La persistance est une conséquence directe de l'autonomie. Les propriétés de l'accommodation de l'état de croyance à la nouvelle information et de la réalisation des opérations de changement, définies dans AGM, ne sont plus valables.
- *La crédibilité des informateurs* : un agent n'accepte pas toute information provenant de n'importe quel agent de son environnement. Il considère certains agents comme crédibles et d'autres non. Parmi les agents considérés crédibles, certains sont plus crédibles que d'autres. Les jugements de crédibilité sont relatifs : chaque agent a ses propres jugements. Un agent i peut considérer un agent j crédible alors qu'un agent k peut le considérer non-crédible.

3.1 Crédibilité des agents informateurs

En recevant un message qui expose de l'information, un agent ne possède pas toujours les moyens pour la justifier. Il ne dispose pas toujours des ressources nécessaires qui lui permettent de faire le changement de croyances adéquat par rapport à cette nouvelle information. Cependant, puisque chaque message reçu est accompagné par l'identificateur de la source qui le signe, cette information peut être de grande utilité. En se basant sur la crédibilité de la source, l'agent qui reçoit un message peut effectuer le changement de croyances qui convient.

Les composantes de la crédibilité La crédibilité n'est pas une qualité intrinsèque à l'agent source, mais elle est plutôt le fait de l'agent récepteur. Certes, c'est l'agent source qui est crédible, mais c'est parce que l'agent récepteur le perçoit comme tel qu'il opère auprès de lui les effets de sa crédibilité. La perception qu'a le récepteur de la source est primordiale pour comprendre les effets des sources. La crédibilité repose essentiellement sur deux facteurs : la compétence et la fiabilité.

La *compétence* réfère à la capacité de la source de dire la vérité. Elle est fondée sur les connaissances que possède le communicateur et sur son expertise. Le récepteur doit donc percevoir chez la source une expertise suffisante qui lui permet de parler en connaissance de cause.

La *fiabilité* réfère, quant à elle, à la volonté de dire la vérité. Elle se rapporte à l'intention que l'on perçoit de la source. On l'assimile souvent à l'honnêteté.

Capacité et volonté de dire la vérité doivent être présentes simultanément chez la source d’information afin qu’elle soit perçue comme crédible. Selon Eagly, Wood et Chaiken [EWC78], le récepteur perçoit chez la source un biais de connaissance s’il la considère incompétente et un biais d’information s’il la considère non fiable. Notons que la crédibilité d’un agent ne s’étend pas à tous les objets, mais seulement à ceux concernés par son expérience. Une source est crédible lorsque ses connaissances ou son expérience sont reliées à l’objet de la communication. Nous considérons que les agents du système peuvent présenter un biais de connaissance mais pas un biais d’information. Tous les agents du système sont considérés honnêtes. Par conséquent, la crédibilité se réduit dans ce cas à la compétence.

Les effets des sources non-crédibles Hovland et Weiss [HW51] figurent parmi les premiers chercheurs en psychologie sociale à avoir analysé empiriquement l’effet des sources crédibles sur les attitudes des récepteurs. Leurs travaux sont considérés comme des références sur la question. En attribuant un même message tantôt à une source crédible, tantôt à une source non-crédible, Hovland et Weiss ont été capables de suivre les changements d’attitude des récepteurs du message selon la source qui le signe. Leurs résultats ont démontré que les sources jugées non-crédibles ont opéré des changements d’attitude négligeables par rapport aux sources crédibles. De ce fait, nous considérons dans cette étude que les agents non-crédibles n’ont aucune influence. L’influence des sources crédibles varie selon un jugement comparatif de la crédibilité.

4 Cadre théorique

Dans cette section, nous définissons un cadre théorique simple qui nous permet d’exprimer les différentes considérations définies dans la section précédente. Ce cadre théorique s’inspire de celui utilisé par AGM mais diffère de ce dernier sur plusieurs points. Le but essentiel de ce cadre théorique est de rendre compte de certains aspects du changement des croyances dans un SMA en utilisant la logique étiquetée LDS.

Pour nous, un agent a des capacités de raisonnement limitées, autrement dit, il est non-omniscient. Pour y arriver à de telles capacités limitées, nous avons développé une approche similaire à Fagin et al.[FHMV95]. Dans cette approche, un agent i est dit connaître explicitement p (noté $X_i p$), s’il sait implicitement p (i.e., $K_i p$), tout en étant “conscient” de p (noté $A_i p$). Formellement, cela peut s’écrire :

$$X_i p \stackrel{\text{déf}}{=} K_i p \wedge A_i p.$$

Pour nous, dire que i est conscient de p , revient à dire qu’il est capable de justifier la vérité de p en un temps fini. Ceci nous amène donc à :

$$A_i \phi \equiv \exists \{a_1, a_1, \dots, a_n\} \text{ tel que } \{a_1, a_1, \dots, a_n\} : \phi \wedge \tau_i[\{a_1, a_1, \dots, a_n\} : \phi].$$

Autrement dit, i est conscient de ϕ ssi il existe des justifications a_1, a_1, \dots, a_n pour ϕ et que i a la capacité de les utiliser en un temps fini pour justifier ou supporter ϕ . L'expression, $\tau_i[\{a_1, a_1, \dots, a_n\} : \phi]$ (qui peut être vraie ou fausse) reflète donc (quand elle est vraie) cette capacité d'utiliser les justifications a_1, a_1, \dots, a_n en un temps fini.

On pourra dès lors soit écrire (ceci découle directement de la définition) $X_i p$ pour signifier que l'agent i croit explicitement p et que donc, il a les moyens de la justifier en un temps fini. Soit écrire : $X_i(\{a_1, a_1, \dots, a_n\} : \phi)$ pour signifier que l'agent i croit explicitement ϕ et qu'il peut le justifier en utilisant $\{a_1, a_1, \dots, a_n\}$, et ce en un un temps fini. La différence entre les deux réside dans le fait que le processus de recherche d'arguments n'est pas nécessaire dans le deuxième cas, alors qu'il l'est dans le premier cas.

Si on pose H l'ensemble des justification du genre $\{a_1, a_1, \dots, a_n\}$, on pourra alors écrire :

$$X_i(H : \phi) \rightarrow X_i \phi.$$

$$X_i(H : \phi) \rightarrow K_i(H : \phi).$$

La deuxième expression découle du fait que toute croyance explicite implique une croyance implicite (découle de la définition).

Il reste maintenant à savoir s'il faut écrire $K_i(H : \phi)$ ou tout simplement $K_i(\phi)$ et que veut-on exprimer par l'une et l'autre des deux expressions ? En fait les deux expressions sont permises au niveau de notre formalisme. Ainsi :

- $K_i(H : \phi)$ signifie que i sait implicitement comment justifier ϕ sans savoir explicitement comment le faire. Tel est le cas par exemple si H justifie ϕ par l'utilisation de la clé X45 (pour la décrypter par exemple). Dans ce cas, si i ne sait pas *comment* utiliser une telle clé, sa connaissance va donc rester au niveau de l'implicite.
- $K_i(\phi)$ signifie cette fois-ci que i sait que ϕ est vraie, sans qu'il sache en fait comment le justifier. Ceci paraît assez conforme à une connaissance implicite.

Le lecteur intéressé par les détails de cette approche peut se référer à la thèse de Jarras [Jar01].

En utilisant notre modèle et comme pour AGM, un agent peut avoir trois attitudes de croyance par rapport à une proposition :

- *Acceptation* : $K_i p$ ou $K_i(H : p)$. Rappelons que la première formule précise que i croit implicitement p , alors que la deuxième signifie que i croit implicitement que l'ensemble de justifications H supporte p . Toutefois, i n'a pas la capacité de déterminer p en utilisant H .
- *Rejet* : $K_i \neg p$ ou $K_i(H : \neg p)$. Les explications sont similaires à celles du précédent item où il suffit de remplacer p par $\neg p$.
- *Indétermination* : $\neg K_i p \wedge \neg K_i \neg p$.

L'acceptation, la non-acceptation et l'indétermination ont respectivement les sens de croyance implicite, croyance implicite de la négation et ignorance.

4.1 Les entrées

Contrairement à AGM, où aucune distinction entre les types d'entrées n'est considérée, nous distinguons deux principaux types d'entrées :

- *Les entrées internes* : ces entrées correspondent aux résultats obtenus par l'agent de sa propre "expérience". Le contact entre l'agent et son environnement est direct. Par exemple, le résultat d'une perception est une entrée interne.
- *Les entrées externes* : ces entrées correspondent aux informations reçues des autres agents du système. Dans ce cas, le contact entre l'agent et son environnement est indirect : il se fait à travers les autres agents.

Cette distinction est primordiale car elle nous permet de différencier deux types de croyances :

- Les croyances privées : croyances acquises par l'agent à partir des entrées internes.
- Les croyances communiquées : croyances acquises à partir des entrées externes.

Pour rendre explicite la différence entre croyances privées et croyances communiquées, il est nécessaire d'associer à toute croyance une information additionnelle précisant sa provenance (l'identité de l'agent source). Dans ce cas on pourrait utiliser l'étiquette pour identifier la source. Ainsi, on pourrait par exemple écrire :

$$K_i(H : p) \text{ avec } H = \{\dots, Inform(i, j, p) \wedge Sincere(j), \dots\}$$

pour préciser que l'agent i connaît p parce que j l'a informé et que celui-ci est considéré par i comme étant sincère. Ainsi, i a gardé trace de son informateur et du fait qu'il soit sincère.

On aurait pu également gérer les sources par des indices comme suggéré par Djefal [Dje96, DCd98]. Ce genre de gestion pose toutefois, quelques problèmes dus principalement à la lourdeur induite par la gestion des indices d'une part, et du fait, qu'il ne rend pas compte des justifications que les agents pourraient garder comme traces, d'autre part.

Le fait de garder trace des croyances au niveau des étiquettes permet, entre autres, de :

- *choisir la stratégie de changement de croyances adéquate*. Pour expliquer ce point considérons le cas suivant : un agent i reçoit p de j considéré crédible en p . Dans ce cas, i pourrait par exemple faire une expansion par p tout en incluant au niveau de ses justifications, la source de p , ce qui peut la caractériser, ainsi que les justifications soutenant p que cette source lui a fourni. Après cela, le même i reçoit $\neg p$ d'un autre agent k . Dans ce cas, l'agent i aura à évaluer les deux sources j et k et voir lequel des deux est plus crédible. Ainsi par exemple, si j est plus crédible que k , cela veut dire que i va opter pour p et garder les justifications provenant de j . Au contraire, si i juge k plus crédible en p que j alors il pourrait mettre à jour ses connaissances par $\neg p$ et lui associer toutes les justifications qu'il reçoit de k .

- *propager les effets de la remise en cause de la crédibilité d'un agent.* En effet, l'agent récepteur peut changer d'avis, pour une raison ou une autre, sur la crédibilité d'un autre agent de qui, des informations ont été, auparavant, acceptées. Ainsi, un agent considéré, auparavant, comme crédible peut cesser de l'être. Dans ce cas, les informations provenant de lui doivent être révisées. Pour ce faire, l'agent récepteur doit être capable de les identifier. La propagation peut aussi avoir lieu dans le cas où l'agent informateur revient, pour une raison ou une autre, sur les informations qu'il a précédemment communiquées.

4.2 Changements d'attitudes et opérations de changement

En prenant en considération le fait qu'un agent peut persister dans ses attitudes de croyances lors de ses interactions avec d'autres agents, de nouveaux changements d'attitudes de croyance sont ajoutés à ceux définis par AGM. Ces derniers correspondent à l'opération de persistance. Le tableau 1 donne la liste complète de ces attitudes et des opérations correspondantes.

Opération	attitude initiale	attitude finale	Commentaires
Expansion	$\neg K_i p \wedge \neg K_i \neg p$	$K_i(H : p)$	(a)
	$\neg K_i p \wedge \neg K_i \neg p$	$K_i(H : \neg p)$	
Révision	$K_i \neg p$ ou $K_i(H : \neg p)$	$K_i(H' : p)$	(b)
	$K_i p$ ou $K_i(H : p)$	$K_i(H' : \neg p)$	
Contraction	$K_i p$ ou $K_i(H : p)$	$\neg K_i p \wedge \neg X_i \neg p$	(c)
	$K_i \neg p$ ou $K_i(H : \neg p)$	$\neg K_i p \wedge \neg X_i \neg p$	
Persistance	$K_i p$ ou $K_i(H : p)$	$K_i(H' : p)$	(d)
	$K_i \neg p$ ou $K_i(H : \neg p)$	$K_i(H' : \neg p)$	
	$\neg K_i p \wedge \neg K_i \neg p$	$\neg K_i p \wedge \neg K_i \neg p$	

TABLE 1. Changements d'attitudes et opérations de changement.

Pour les changements représentés ci-dessus, nous pouvons faire les commentaires suivants :

- Pour le cas (a), l'attitude de i passe de l'indétermination à l'acceptation ou au rejet. Dans l'un et l'autre des cas, i garde trace du changement dans l'ensemble des justifications H .
- Pour (b), i part d'une attitude initiale qui pourrait être justifiable (par H par exemple) ou non, et effectue une révision. Une fois celle-ci effectuée, i se trouve avec une acceptation ou un rejet justifiable par H' qui peut être considéré comme une mise à jour de la justification de départ (que cette justification soit H ou \emptyset).
- Pour le cas (c) i part d'une acceptation ou d'un rejet justifiable, ou non, et se trouve dans une position d'ignorance où nous n'avons pas senti le besoin de "justifier l'ignorance", dans la mesure où (généralement) on n'a pas besoin de justifier ce qu'on ignore.

- Le cas (d) est similaire aux cas précédents et de ce fait il se comprend aisément.

4.3 Opérateurs de crédibilité

Le concept de crédibilité est utilisé comme facteur externe intervenant dans le changement des croyances. Pour en tenir compte, nous introduisons les opérateurs suivants :

- $C_{i,p}^j$: i considère j crédible en p ;
- $\neg C_{i,p}^j$: i considère j non-crédible en p .

On supposera ici que les opérateurs du type $C_{i,p}^x$ peuvent être soit comparables entre eux sur une échelle numérique, soit incomparables entre eux. Dans le cas où $C_{i,p}^j > C_{i,p}^k$ on notera cette relation par $SC_{i,p}^{j,k}$ afin de préciser que l'agent i considère j plus crédible que k en p . À l'inverse et dans le cas où $C_{i,p}^k > C_{i,p}^j$, la relation est notée $SC_{i,p}^{j,k}$ pour signifier cette fois-ci que l'agent i considère j plus crédible que k en p . Si jamais, $C_{i,p}^k = C_{i,p}^j$, on note cela par $EC_{i,p}^{j,k}$ afin de préciser que i considère, cette fois-ci, j aussi crédible que k en p . Finalement, il se peut que $C_{i,p}^j$ et $C_{i,p}^k$ soient incomparables entre eux, on notera alors cela par $IC_{i,p}^{j,k}$ afin de préciser que l'agent i considère j incomparable avec k en terme de crédibilité en p .

Si on considère $C_{i,p}^*$ l'ensemble de tous les agents considérés par i comme crédibles en p , $C_{i,p}^* = \{j \mid C_{i,p}^j\}$ alors $SC_{i,p}$, $EC_{i,p}$ et $IC_{i,p}$ définissent une structure de préordre partiel sur $C_{i,p}^*$. Cela signifie qu'en terme de crédibilité en p , l'agent i classe certains éléments de $C_{i,p}^*$ du "meilleur" au "moins bon" avec d'éventuels ex æquo.

À ces opérateurs de crédibilité, nous associons les propriétés suivantes :

Propriétés des opérateurs de crédibilité

- $C_{i,p}^j \equiv C_{i,\neg p}^j$
Si i considère j crédible en p , il le considère, aussi, crédible en $\neg p$. L'inverse, aussi, est vraie.
- $SC_{i,p}^{j,k} \vee EC_{i,p}^{j,k} \vee IC_{i,p}^{j,k} \rightarrow C_{i,p}^j \wedge C_{i,p}^k$
La comparaison des agents en terme de crédibilité ne s'applique que pour les agents considérés crédibles.
- $SC_{i,p}^{j,k} \rightarrow \neg SC_{i,p}^{k,j}$
Si i considère j plus crédible en p que k , i ne peut considérer k plus crédible en p que j .
- $SC_{i,p}^{j,k} \wedge SC_{i,p}^{k,l} \rightarrow SC_{i,p}^{j,l}$
Si i considère j plus crédible en p que k et ce dernier plus crédible en p que l alors i considère j plus crédible en p que l .
- $EC_{i,p}^{j,k} \equiv EC_{i,p}^{k,j}$
Si i considère j aussi crédible en p que k alors i considère k aussi crédible en p que j . L'inverse, aussi, est vraie.

- $EC_{i,p}^{j,k} \wedge EC_{i,p}^{k,l} \rightarrow EC_{i,p}^{j,l}$
Si i considère j aussi crédible en p que k et ce dernier (k) aussi crédible en p que l alors i considère j aussi crédible en p que l .
- $IC_{i,p}^{j,k} \equiv IC_{i,p}^{k,j}$
Si i considère j incomparable avec k en terme de crédibilité en p alors i considère k incomparable avec j en terme de crédibilité en p . L'inverse, aussi, est vraie.
- $SC_{i,p}^{j,k} \wedge EC_{i,p}^{k,l} \rightarrow SC_{i,p}^{j,l}$
Si i considère j plus crédible en p que k et ce dernier (k) aussi crédible en p que l alors i considère j plus crédible en p que l .
- $EC_{i,p}^{j,k} \wedge SC_{i,p}^{k,l} \rightarrow SC_{i,p}^{j,l}$
Si i considère j aussi crédible en p que k et k plus crédible en p que l alors i considère j plus crédible en p que l .

4.4 Communication entre agents

Dans cet article, nous utilisons l'approche communication la plus couramment utilisée : l'approche linguistique basée sur les actes du langage naturel [Sea69, Van90]. Il faut dire que de tels actes sont basés sur les états mentaux des intervenants (croyances, désirs, intentions, etc.) et de ce fait, ils sont donc cohérents avec nos agents considérés comme systèmes logiques.

Nous considérons ici, que la communication se fait point à point. Cependant, un agent peut recevoir plusieurs messages en même temps. Un message est défini par l'agent informateur (source), l'agent récepteur et l'information convoyée par le message. Nous utiliserons, pour décrire l'envoi de messages, l'opérateur $Envoie(i, j, m)$ qui a pour signification "l'agent i envoie à l'agent j le message m ". m est du type $\langle f, p \rangle$, c'est à dire une force illocutoire f et un contenu propositionnel p , comme le spécifie Searle et Vanderveken. Par exemple, $\langle Informe_{ij}, p \rangle$ signifie que i informe j que p , en utilisant donc la force illocutoire d'un assertif et une proposition p .

$Envoie(i, j, m)$ est satisfait si le message m dont le contenu est $\langle f, p \rangle$ a bien été envoyé par i à j . Du fait que l'accent est mis sur le récepteur du message puisque c'est lui qui est concerné par le changement de croyances, nous utiliserons l'opérateur $Reçoit(j, i, m)$ qui signifie "l'agent j reçoit de l'agent i le message dont le contenu est m ". $Reçoit(j, i, m)$ est satisfait si le message, dont le contenu est m et l'agent émetteur est i , a bien été reçu par l'agent j .

Nous supposons pour finir, que tout agent participant à la communication est *sincère*. Ses énoncés reflètent donc fidèlement son état mental. Ainsi si un agent i dit que «la Juventus a gagné», alors i croit effectivement que «la Juventus a gagné». Cette hypothèse est destinée essentiellement à rejeter le mensonge. De ce fait, un agent dans notre approche, ne raisonne pas sur la sincérité pour adopter ou rejeter une nouvelle information comme dans d'autres approches [CL90], mais plutôt sur le fait de savoir si l'émetteur du message est fiable ou non ? La sincérité des agents nous semble difficilement appréhendable dans la mesure où lorsqu'on est confronté à un agent décidé à mentir, et que nous ne supposons pas a priori qu'il est non-sincère, il nous est relativement difficile de prouver qu'il

ment. Chez les humains, ce n'est que par recoupement d'informations, analyse du visage et des émotions reflétés, qu'on arrive (parfois) à détecter le mensonge. Autant de considérations qui sont en dehors du présent travail. Dans le cadre de notre travail et de la logique étiquetée, on pourrait utiliser les justifications pour détecter le mensonge. Pour cela, on pourrait demander à l'agent informateur de donner les raisons qui sous-tendent son énoncé et de vérifier les dites raisons, mais cela aussi n'est pas toujours facile à faire dans la mesure où l'agent informateur pourrait continuer à ne pas être sincère et s'arranger pour fournir les «bonnes raisons» même si elles ne sont pas vraies.

5 Stratégies de changement des croyances

En tenant compte de la crédibilité, nous allons aborder la révision des croyances à partir d'entrées externes en termes de situations. Pour chaque situation, un ensemble de stratégies de changement de croyances est proposé. Cependant, seul le changement d'attitude par rapport à la proposition concernée par le changement est défini. Toutes les stratégies sont définies par la forme générale suivante :

$$\text{attitude-}i \wedge \text{entrée-externe} \wedge \text{crédibilité} \xrightarrow{\text{opération}} \text{attitude-}f$$

- *attitude- i* : attitude de croyance initiale de l'agent i par rapport à p ;
- *entrée-externe* : information(s) reçue(s) des autres agents ;
- *crédibilité* : jugement de crédibilité de l'agent(s) informateur(s) ;
- *opération* : opération de changement de croyances à appliquer ;
- *attitude- f* : attitude de croyance de l'agent i par rapport à p après le changement de croyances.

Dans ce qui suit et afin de ne pas alourdir nos formules, nous simplifierons l'écriture de $Reçoit(i, j, \langle Informe_{ji}, p \rangle)$ en l'a ramenant à simplement : $Informe(j, i, p)$. Dans le cas où j informe i que p tout en le justifiant par H par exemple, la formule que nous adopterons, sera cette fois-ci : $Informe(j, i, H : p)$

5.1 Les croyances privées

Situation 1 L'agent récepteur i a une attitude indéterminée par rapport à p et un agent j l'informe que p . Deux stratégies sont envisageables pour cette situation :

- *Persistence1* : si l'agent i considère l'agent j non-crédible en p , l'agent i persiste dans son attitude indéterminée.

$$\neg K_i p \wedge \neg K_i \neg p \wedge Informe(j, i, p) \wedge \neg C_{i,p}^j \xrightarrow{\text{Persistence}} \neg K_i p \wedge \neg K_i \neg p$$

- *Expansion1* : si l'agent i considère l'agent crédible en p , l'agent i accepte p de j .

$$\neg K_i p \wedge \neg K_i \neg p \wedge Informe(j, i, p) \wedge C_{i,p}^j \xrightarrow{\text{Expansion}} K_i(H : p)$$

Comme on peut le constater, i accepte p certes, mais il lui adjoint des justifications pour tenir compte particulièrement de : (i) la source, c.-à-d. $Informe(i, j, p)$ et (ii) la *la raison* qui l'a amené à changer de croyances, soit $C_{i,p}^j$. Une telle expansion, mènera i à croire explicitement à p (soit $X_i(p)$) si $\tau_i(H : p)$

Situation 2 L'agent récepteur i a une attitude indéterminée par rapport à p et un agent j l'informe que $H : p$. Deux stratégies sont envisageables pour cette situation :

- *Persistence* : si l'agent i considère l'agent j non-crédible en p , l'agent i persiste dans son attitude indéterminée.

$$\neg K_i p \wedge \neg K_i \neg p \wedge Informe(j, i, H : p) \wedge \neg C_{i,p}^j \xrightarrow{\text{Persistence}} \neg K_i p \wedge \neg K_i \neg p$$

On pourrait argumenter ici que si i a les moyens de vérifier que les justifications H supportent bien p , il pourrait accepter ce que j suggère et remettre en question sa connaissance sur la crédibilité de j . En fait, i ne pourrait faire cela, car il n'a aucun moyen (dans notre formalisme) de vérifier que les justifications avancées par j soient des justifications correctes. Autant alors, qu'il persiste dans son ignorance, quitte à trouver un autre agent plus crédible que j et vérifier auprès de lui.

- *Expansion* : si l'agent i considère l'agent j crédible en p , l'agent i accepte $H : p$ de j .

$$\neg K_i p \wedge \neg K_i \neg p \wedge Informe(j, i, H_1 : p) \wedge C_{i,p}^j \xrightarrow{\text{Expansion}} K_i(H_2 : p)$$

Comme on peut le constater, i accepte p certes, mais il met aussi à jour ses justifications qui passent de H_1 à H_2 . Dans H_2 , on devrait trouver : (i) H_1 , c.-à-d. les justifications transmises par le crédible agent j , (ii) $Informe(j, i, H_1 : p)$ qui représente la source d'où est venue p et, (iii) $C_{i,p}^j$ qui représente *la raison* qui a amené i à changer de croyances.

Partant de cette expansion, l'agent pourrait vérifier si $\tau_i(H_2 : p)$ (C'est à dire s'il a la capacité de justifier p en utilisant H_2 , et ce, en un temps fini.).

Si c'est le cas alors $X_i(H_2 : p)$ et par conséquent aussi $X_i(p)$.

Il convient de remarquer que les situations 1 et 2 peuvent aisément s'écrire sous forme condensée, en raisonnant sur les croyances justifiées ayant la forme $H : p$. En effet, la forme $K_i p$ n'est qu'un cas particulier de $K_i(H : p)$ où H est un ensemble vide (\emptyset) de justifications. Dès lors, les situations ci-dessous vont porter sur des croyances justifiées ayant la forme générale $H : p$ et les sous situations sans justifications peuvent aisément se déduire en remplaçant H par \emptyset .

Situation 3 L'agent récepteur i croit $H_1 : \neg p$ et un agent j lui communique $H_2 : p$. Pour ce type de situation, on a également deux stratégies :

- *Révision* : si l'agent i considère l'agent j plus crédible en p que lui, il passe de la non-acceptation de p à l'acceptation de p de j , en changeant aussi

ses justifications.

$$K_i(H_1 : \neg p) \wedge Informe(j, i, H_2 : p) \wedge SC_{i,p}^{j,i} \xrightarrow{\text{Révision}} K_i(H_3 : p)$$

avec $H_3 = \{Informe(j, i, H_1 : p) \wedge SC_{i,p}^{j,i}\} \sqcup H_2$

- *Persistence* : si l'agent i considère l'agent j non-crédible en p ou pas plus crédible que lui, alors l'agent i persiste dans son attitude de non-acceptation de p avec les mêmes justifications.

$$K_i(H_1 : \neg p) \wedge Informe(j, i, H_2 : p) \wedge \neg C_{i,p}^j \vee \neg SC_{i,p}^{j,i} \xrightarrow{\text{Persistence}} K_i(H_1 : \neg p)$$

Situation 4 Cette situation correspond au cas où l'agent récepteur i a une attitude de croyance indéterminée par rapport à p , un agent j lui communique $H_1 : p$ et un autre agent k lui communique $H_2 : \neg p$. Pour ce type de situation, on a trois stratégies :

- *Expansion1* : si l'agent i considère l'agent j crédible en p et l'agent k non-crédible en p ou s'il considère j plus crédible en p que k , alors i accepte l'information de j et passe de l'indétermination à l'acceptation de p avec pour justifications $H_3 = \{Informe(j, i, H_1 : p) \wedge C_{i,p}^j \wedge \neg C_{i,p}^k \vee SC_{i,p}^{j,k}\} \sqcup H_1$.

$$\neg K_{i,p} \wedge \neg K_{i,\neg p} \wedge Informe(j, i, H_1 : p) \wedge Informe(k, i, H_2 : \neg p) \wedge ((C_{i,p}^j \wedge \neg C_{i,p}^k) \vee SC_{i,p}^{j,k}) \xrightarrow{\text{Expansion}} K_i(H_3 : p)$$

- *Expansion2* : si l'agent i considère l'agent k crédible en p et l'agent j non-crédible en p ou s'il considère k plus crédible en p que j , alors i accepte l'information de k et passe de l'indétermination à la non-acceptation de p avec pour justifications $H_3 = \{Informe(k, i, H_2 : \neg p) \wedge C_{i,p}^k \wedge \neg C_{i,p}^j \vee SC_{i,p}^{k,j}\} \sqcup H_2$.

$$\neg K_{i,p} \wedge \neg K_{i,\neg p} \wedge Informe(j, i, H_1 : p) \wedge Informe(k, i, H_2 : \neg p) \wedge ((C_{i,p}^k \wedge \neg C_{i,p}^j) \vee SC_{i,p}^{k,j}) \xrightarrow{\text{Expansion}} K_i(H_3 : \neg p)$$

- *Persistence* : si l'agent i considère les agents j et k non-crédibles en p ou aussi crédibles l'un que l'autre ou s'il ne peut les comparer en termes de crédibilité en p alors il persiste dans son indétermination par rapport à p .

$$\neg K_{i,p} \wedge \neg K_{i,\neg p} \wedge Informe(j, i, H_1 : p) \wedge Informe(k, i, H_2 : \neg p) \wedge ((\neg C_{i,p}^j \wedge \neg C_{i,p}^k) \vee EC_{i,p}^{j,k} \vee IC_{i,p}^{j,k}) \xrightarrow{\text{Persistence}} (\neg K_{i,p} \wedge \neg K_{i,\neg p})$$

5.2 Les croyances communiquées

Situation 1 L'agent i , après avoir accepté une information de la part d'un agent j , reçoit la même information de la part d'un autre agent k . Cette situation est considérée pour montrer que les opérations de changement des croyances ne sont pas idempotentes lorsqu'il s'agit des croyances communiquées. Pour ce type de situation, on a trois stratégies :

- *Persistence* : si l'agent i considère l'agent k non-crédible en p ou moins crédible que l'agent j ou incomparable en terme de crédibilité avec j alors l'agent i persiste dans son attitude d'acceptation de p de j .

$$K_i(H_1 : p) \wedge \text{Informe}(k, i, H_2 : p) \wedge (\neg C_{i,p}^k \vee SC_{i,p}^{j,k} \vee IC_{i,p}^{j,k}) \xrightarrow{\text{Persistence}} K_i(H_1 : p)$$

Il faudra retenir ici que H_1 contient au départ $\text{Informe}(j, i, H' : p) \sqcup H'$ (avec H' une quelconque justification donnée par j au départ), et reflète donc la provenance de p , c.-à-d. l'agent j . Une fois que l'agent k informe i , celui-ci persiste à faire confiance à j puisqu'il ne change pas les justifications qui sous-tendent p .

- *Expansion1* : si l'agent i considère k plus crédible en p que j alors i passe de l'acceptation de p de j à l'acceptation de p de k .

$$K_i(H_1 : p) \wedge \text{Informe}(k, i, H_2 : p) \wedge SC_{i,p}^{j,k} \xrightarrow{\text{Expansion}} K_i(H_3 : p)$$

À noter que i part avec une justification H_1 qui contient donc $\text{Informe}(j, i, H' : p) \sqcup H'$ (avec H' une quelconque justification donnée par j au départ), et reflète donc la provenance de p , c.-à-d. l'agent j . À l'arrivée, i va former un nouveau ensemble de justifications (H_3) basée essentiellement sur H_2 et $\text{Informe}(k, i, H_2 : p)$, pour préciser qu'il passe de l'acceptation de p de j à l'acceptation de p de k .

- *Expansion2* : si l'agent i considère k aussi crédible que j en p , alors i passe de l'acceptation de p de j à l'acceptation de p de j et k .

$$K_i(H_1 : p) \wedge \text{Informe}(k, i, H_2 : p) \wedge EC_{i,p}^{j,k} \xrightarrow{\text{Expansion}} K_i(H_3 : p)$$

Cette fois-ci, au départ H_1 reflète $\text{Informe}(j, i, H' : p) \sqcup H'$ et à l'arrivée, H_3 reflète $\text{Informe}(k, i, H_2 : \neg p) \sqcup H_1 \sqcup H_2$. De cette façon, l'agent i est passé de l'acceptation de p de j à une acceptation de p de j et de k .

Situation 2 L'agent i reçoit d'un agent k des informations qui contredisent celles acceptées, auparavant, de la part d'un autre agent j . Pour ce type de situation, on a trois stratégies :

- *Persistence* : si l'agent i considère l'agent k non-crédible ou s'il considère j plus crédible en p que k ou incomparable en termes de crédibilité avec ce dernier (l'agent k) alors l'agent i persiste dans son attitude d'acceptation

de p de j .

$$K_i(H_1 : p) \wedge \text{Informe}(k, i, H_2 : \neg p) \wedge (\neg C_{i,p}^k \vee SC_{i,p}^{j,k} \vee IC_{i,p}^{j,k}) \xrightarrow{\text{Persistence}} K_i(H_1 : p)$$

Ici aussi i part avec une justification H_1 qui contient $\text{Informe}(j, i, H' : p) \sqcup H'$ (avec H' une quelconque justification donnée par j au départ), et reflète donc la provenance de p , c.-à-d. l'agent j . Une fois que l'agent k informe i , celui-ci persiste à faire confiance à j puisqu'il ne change pas les justifications qui sous-tendent p .

- *Contraction* : si l'agent i considère l'agent k aussi crédible en p que j , alors l'agent i passe de l'acceptation de p de j à l'indétermination par rapport à p .

$$K_i(H_1 : p) \wedge \text{Informe}(k, i, H_2 : \neg p) \wedge EC_{i,p}^{k,j} \xrightarrow{\text{Contraction}} \neg K_i p \wedge \neg K_i \neg p$$

- *Révision* : si l'agent i considère l'agent k plus crédible en p que j , alors l'agent i passe de l'acceptation de p de j à l'acceptation de $\neg p$ de k .

$$K_i(H_1 : p) \wedge \text{Informe}(k, i, H_2 : \neg p) \wedge SC_{i,p}^{k,j} \xrightarrow{\text{Révision}} K_i(H_3 : \neg p)$$

Comme on peut le constater, au départ H_1 reflète $\text{Informe}(j, i, H' : p) \sqcup H'$ et à l'arrivée, H_3 reflète $\text{Informe}(k, i, H_2 : \neg p) \sqcup H_2$. De cette façon, l'agent i est passé de l'acceptation de p de j à une acceptation de $\neg p$ de k en changeant également de justifications qui proviennent cette fois-ci de k et non de j .

6 Conclusion et perspectives

Partant du constat que la révision classique ne permettait pas de traiter la révision des croyances au niveau d'un système multiagent, nous avons expliqué que, dans de tels systèmes où les agents communiquent entre eux et ont le contrôle total sur leurs états de croyance respectifs, d'autres considérations doivent être prises en compte, en particulier : (i) un agent peut persister sur ses attitudes de croyance; (ii) un agent garde trace des sources de ses croyances.

À partir de ces considérations, nous avons fait la distinction entre entrées internes et entrées externes. Ceci nous a permis de définir deux types de croyances : les croyances privées et les croyances communiquées. Forts de cette typologie, nous avons étudié le changement au niveau de ces deux types de croyances à partir de la communication entre agents. Plusieurs situations ont été considérées, et pour chaque situation, plusieurs stratégies de changement des croyances basées sur la crédibilité des communicateurs ont été proposées. Ces stratégies reposent spécialement sur les points suivants : (1) un agent n'accepte d'informations que des agents qu'il considère crédibles; (2) un agent ne change d'attitude de croyance que s'il considère l'agent qui lui communique l'information inconsistante plus crédible que lui; (3) en cas de contradiction entre deux sources

considérées incomparables ou égales en crédibilité, un agent persiste sur son attitude de croyance si elle est privée (dans le cas contraire, il passe à l'indétermination); (4) pour toute croyance, un agent garde trace de la source la plus crédible.

Comme perspectives de continuité, nous proposons (1) d'étudier plus en détail la révision des croyances privées, et ce en accord avec les notions d'importance (de chacune des croyances), de préférence (entre les croyances), de pertinence (des croyances) ou d'enracinement épistémique; (2) de faire le lien entre la notion de crédibilité et l'importance, la pertinence, la préférence ou l'enracinement épistémique; (3) de considérer le cas des croyances obtenues par combinaison des croyances privées et des croyances communiquées; (4) de définir les critères qui permettent à un agent de choisir entre les différents types de contraction et de révision que nous avons introduits dans ce papier.

Références

- [AGM85] C. E. Alchourròn, P. Gärdenfors, and D. Makinson. On the logic of theory change : partial meet functions for contraction and revision. *Journal of Symbolic Logic*, 50 :510–530, 1985.
- [AM85] C. E. Alchourròn and D. Makinson. On the logic of theory change : Safe contraction. *Studia Logica*, 44 :405–422, 1985.
- [CL90] P. R. Cohen and H. J. Levesque. Intention is choice with commitment. *Artificial Intelligence*, 42 :213–261, 1990.
- [DCd98] S. Djeflal and B. Chaib-draa. Stratégie de croyances basée sur la crédibilité des agents. *Revue d'intelligence artificielle*, 12 :103–123, 1998.
- [Dje96] S. Djeflal. Dynamique des croyances : aspects multi-agents. Master's thesis, Département d'informatique, Université Laval, Québec, 1996.
- [EWC78] A.H. Eagly, W. Wood, and S. Chaiken. Causal inferences about communicators and their effect on opinion change. *Journal of Personality and Social Psychology*, 36 :161–202, 1978.
- [FHMV95] R. Fagin, I. Y. Halpern, Y. Moses, and M. Y. Vardi. *Reasoning about Knowledge*. The MIT Press : Cambridge, MA, 1995.
- [Gab96] D. M. Gabbay. *Labelled Deductive Systems - Volume I*. Clarendon Press, Oxford, 1996.
- [Gär88] P. Gärdenfors. *Knowledge in Flux*. The MIT Press, 1988.
- [HW51] C.I. Hovland and W. Weiss. The influence of source credibility on communication influence. *The Public Opinion Quarterly*, 15 :635–650, 1951.
- [Jar01] I. Jarras. *Raisonnement d'agents considérés comme des systèmes logiques : une approche basée sur les étiquettes*. PhD thesis, Département d'informatique, Université Laval, 2001.
- [Sea69] J. R. Searle. *Speech Acts : An Essay in the Philosophy of Language*. Cambridge University Press : Cambridge, England, 1969.
- [Van90] D. Vanderveken. On the unification of speech act theory and formal semantics. In P. R. Cohen, J. Morgan, and M. E. Pollack, editors, *Intentions in Communication*, pages 195–220. The MIT Press : Cambridge, MA, 1990.