

De l'intercompréhension à la complémentarité des relations sociales humain-machine

Florian Pecune, Magalie Ochs, Catherine Pelachaud

CNRS-LTCl, Télécom ParisTech

{pecune, ochs, pelachaud}@telecom-paristech.fr

Résumé

Dans cet article, nous décrivons un modèle conceptuel des attitudes sociales entre un agent conversationnel animé et un humain. Ce modèle propose une représentation des attitudes sociales de l'agent et de l'utilisateur considérant l'influence à la fois du contexte social et des caractéristiques propres à l'agent. L'intercompréhension sociale humain-machine n'étant pas suffisante pour la création d'une relation à long-terme, elle peut être complétée par le biais de la théorie de la complémentarité, à travers la théorie de la congruence. Il y a congruence lorsque l'attitude sociale de l'utilisateur est conforme aux attentes de l'agent suivant sa propre attitude vis-à-vis de la situation.

Mots clés : *agent conversationnel – attitude sociale – théorie de la congruence – théorie de la complémentarité*

Introduction

Les machines occupent de nos jours une place prépondérante dans notre quotidien. Les études présentées par (Krämer, 2008) et (Reeves & Nass, 1996) montrent ainsi que les règles sociales régissant les interactions humain-machine peuvent être similaires à celles entre deux humains. Si le rêve de Bledsoe (Bledsoe, 1986) de voir des ordinateurs créer des liens avec des individus est encore loin d'être une réalité, de nombreux efforts sont cependant réalisés pour attribuer à ces machines qui nous entourent une apparence et des propriétés typiquement humaines.

Outre des robots tels que Nao, capable de parler et de reconnaître de simples commandes vocales, il existe des agents conversationnels animés (ACAs), créés par l'ordinateur et ayant la capacité de communiquer avec l'humain au travers un langage verbal et non-verbal (Cassell, 2000). Ces agents incarnent des rôles divers et variés allant de guide de musée (Swartout & al, 2010) à coach sportif (Bickmore & Picard, 2005). Cependant, très peu de ces agents sont dotés d'une modélisation à la fois de leur propre relation sociale et de celles de l'utilisateur. Or, une telle modélisation peut permettre une intercompréhension sociale entre l'agent et l'utilisateur.

Afin de remplir au mieux ces rôles typiquement incarnés par des êtres humains, il est indispensable que les protagonistes, humains et machines, se comprennent l'un l'autre. Cette intercompréhension (Grandgeorge & MIAC, 2011) s'opère non seulement via la parole, mais également à travers le comportement non verbal, permettant ainsi de transmettre des émotions, des intentions communicatives (Ochs, et al., 2011), mais également des attitudes sociales. L'ACA doit par exemple être capable de détecter des signes d'appréciation ou de dominance chez l'utilisateur et d'agir en conséquence. A son tour, l'utilisateur doit pouvoir percevoir de tels indices sociaux chez l'agent.

Dans cet article, nous abordons dans un premier temps le concept d'intercompréhension dans un contexte social et plus particulièrement dans le contexte des relations sociales entre un humain et une machine durant une interaction. Nous présentons ensuite un modèle conceptuel permettant de représenter les relations sociales d'un agent envers l'humain ainsi que les variables d'influences. Pour finir, nous proposons d'explicitier la complémentarité sociale humain-machine au travers de la théorie de la congruence (Osgood & Tannenbaum, 1955), mettant ainsi en lumière la dynamique des relations sociales entre deux individus.

1. L'intercompréhension dans un contexte social

Lors d'une interaction, deux personnes peuvent exprimer des identités sociales différentes, en fonction de la situation et de la relation sociale qui les lie. L'identité sociale se traduit entre autres au travers de l'attitude sociale exprimée par un individu. Scherer définit ces *attitudes sociales* comme un "*affective style that spontaneously develops or is strategically employed in the interaction with a person or a group of persons, coloring the interpersonal exchange in that situation*" (Scherer, 2005). Lors d'une

interaction, cette attitude sera alors perçue par l'interlocuteur qui exprimera sa propre attitude sociale en retour.

Si l'on applique la définition de l'intercompréhension proposée par (Grandgeorge & MIAC, 2011) à l'attitude sociale dans une interaction entre un agent A et un humain B, il y a *intercompréhension* entre A et B si l'attitude sociale exprimée par A est conforme à l'attitude qu'il imagine que B se représente de lui même, et réciproquement. Par exemple, si l'agent A exprime de la dominance et qu'il pense que B le perçoit comme dominant, et réciproquement. Pour que l'intercompréhension soit possible, l'agent virtuel doit donc être capable de se représenter la notion d'attitude sociale. Dans le domaine des ACAs, les dimensions les plus utilisées pour représenter les attitudes sont celles d'appréciation et de dominance. L'axe d'appréciation est défini comme un sentiment général, positif ou négatif, à propos d'une personne, d'un objet ou d'une situation (Moshkina & Arkin, 2003). L'axe de dominance, quant à lui, représente la capacité d'un agent à influencer le comportement d'un autre individu (Prada & Paiva, 2008). Cette influence est par ailleurs caractérisée par les ressources et les stratégies disponibles pour l'agent (Raven, 2008). Ces deux dimensions peuvent être représentées sous la forme de deux axes orthogonaux formant un circomplexe interpersonnel (Wiggins, Trapnell, & Phillips, 1988).

Dans notre modèle, pour représenter les attitudes sociales sur ce circomplexe, nous proposons d'utiliser deux représentations différentes : *l'attitude sociale exprimée* est modélisée par un vecteur, démarrant à l'origine du circomplexe (Gurtman, 2009). La taille de ce vecteur traduit la rigidité interpersonnelle (Tracey, 2005) : plus la taille du vecteur est grande, moins l'individu pourra adapter son comportement aux situations qu'il rencontre. *L'attitude sociale perçue* - qui correspond à une croyance de l'individu - est quant à elle modélisée par un sous-ensemble du circomplexe, prenant ainsi en compte la part d'incertitude relative à cette croyance. Dans la section suivante, nous présentons plus en détail le modèle conceptuel de ces attitudes sociales.

2. Modèle conceptuel des attitudes sociales exprimées par l'agent

Pour qu'un agent exprime des attitudes sociales appropriées à la situation d'interaction, les mécanismes inhérents aux relations sociales et leur évolution au cours des interactions doivent être modélisés. Dans sa définition des attitudes sociales, Scherer dénote

des origines qui peuvent être stratégiques ou spontanées (Scherer, 2005). Nous nous appuyons donc sur cette définition pour modéliser les relations de l'agent envers l'utilisateur par trois vecteurs différents :

- la *relation ressentie* représente le sentiment global de l'agent (ex: l'agent n'aime pas l'utilisateur). Cette relation évolue au gré des interactions avec l'utilisateur. Par exemple, si ce dernier insulte l'agent à plusieurs reprises, il est fort probable que l'agent l'apprécie moins. La relation ressentie de l'agent dépend de plus de caractéristiques propre à ce dernier telle que sa personnalité (Trapnell & Wiggins, 1990).

- la *relation idéale* indique la relation que l'agent souhaiterait idéalement exprimer dans une situation particulière (ex: un enseignant voudra montrer une forte dominance à ses élèves pendant la rentrée des classes). La taille et la direction du vecteur dépendent ici du rôle de l'agent et de l'image qu'il souhaite donner: un enseignant inflexible sera représenté par un long vecteur dominant tandis qu'un autre plus compréhensif sera représenté par un vecteur plus court. On retrouve ainsi la notion de rigidité interpersonnelle aperçue dans la section 1.

- *L'attitude sociale exprimée* découle donc de la combinaison entre la *relation ressentie* et la *relation idéale*. De plus, comme expliqué par (Ballin, Gillies, & Crabtree, 2004), l'humeur de l'agent peut influencer la valeur d'appréciation (ex: un agent de mauvaise humeur aura tendance à se montrer plus hostile). L'état des ressources de l'agent et de l'utilisateur a quant à lui un impact sur la valeur de dominance du vecteur (Raven, 2008).

3. Modèle conceptuel de l'attitude sociale de l'utilisateur

Comme expliqué dans (Marsella, Pynadath, & Read, 2004), l'un des facteurs clés des relations sociales est la croyance que l'on a à propos d'autrui, aussi appelée théorie de l'esprit (Baron-Cohen, Leslie, & Frith, 1985). Par exemple, un agent n'aura pas la même attitude vis-à-vis d'un utilisateur qu'il pense hostile que face à un utilisateur montrant de l'appréciation. Fondé sur cette théorie, nous proposons dans notre modèle une représentation de la relation de l'utilisateur envers l'agent à travers deux relations :

- *l'attitude perçue* représente l'attitude perçue par l'agent à un instant t (ex: l'agent pense que l'utilisateur montre de la dominance). Dans (Gatica-Perez, 2009), l'auteur dresse une liste d'indices permettant à l'agent de déterminer cette attitude. *L'attitude perçue*

peut être différente de l'attitude réellement exprimée par l'utilisateur : l'agent peut par exemple percevoir des indices de dominance durant une interaction, alors que l'utilisateur veut montrer de l'hostilité. Initialement, *l'attitude perçue* est influencée par la première impression dégagée par l'utilisateur (Cafaro, Vilhjálmsson, Bickmore, Heylen, Jóhannsdóttir, & Valgarðsson, 2012).

- la *relation désirée* est la relation que l'agent désire voir exprimée chez l'utilisateur (ex: l'agent veut que l'utilisateur montre de la soumission). Elle découle de la *relation idéale*, en accord avec la théorie de la complémentarité interpersonnelle (Sadler, Ethier, & Woody, 2011). D'après cette théorie, la *relation désirée* est une symétrie de la *relation idéale* selon l'axe d'appréciation : si l'agent veut montrer de la dominance et de l'appréciation, il s'attend à ce que l'utilisateur exprime de la soumission et de l'appréciation.

L'intercompréhension d'une interaction sociale peut se définir comme la compréhension réciproque de la situation d'interaction, i.e. des attitudes sociales exprimées par l'utilisateur et par l'agent. Cependant, l'intercompréhension n'implique pas nécessairement le bon déroulement de l'interaction : deux individus peuvent se comprendre mutuellement sans que l'interaction soit pour autant satisfaisante. La théorie décrite dans (Sadler, Ethier, & Woody, 2011) montre qu'une complémentarité des relations est un élément important pour la création d'un lien durable entre deux individus. La plupart du temps, un individu dominant préférera interagir avec un autre plus soumis, et réciproquement. Pour concevoir des agents capables d'entretenir des relations à long-terme avec l'utilisateur, ces derniers doivent donc considérer la dynamique des relations sociales pour pouvoir adapter leur comportement afin d'atteindre cette complémentarité. Dans la section suivante, nous présentons la dynamique des attitudes, ainsi que les stratégies envisagées en cas de non complémentarité.

4. Dynamique des attitudes

Afin de modéliser la dynamique de ces relations et ainsi la recherche d'intercompréhension au cours de l'interaction, notre modèle s'appuie sur la *Théorie de la congruence* (Osgood & Tannenbaum, 1955) telle que modélisée dans *Avatar Arena* (Rist, Schmitt, Pelachaud, & Bilvi, 2003). Notre modèle repose donc sur un schéma triangulaire entre un agent A, un utilisateur U et une situation S (Figure 1). Les arrêtes du triangle décrivent la relation de A envers U (*relation ressentie*), la relation de A envers la situation S (*relation idéale*) et l'estimation de A sur la relation de U envers S (*relation*

désirée). Par exemple, l'agent apprécie l'utilisateur, veut montrer de la dominance et s'attend donc à ce que l'utilisateur adopte une attitude soumise.

Selon cette théorie, l'interaction est dite *congruente* quand l'*attitude perçue* de U à propos de S est en adéquation avec la *relation désirée* (Figure 1). Si l'on reprend l'exemple précédent, il y a donc congruence si l'utilisateur montre de la soumission. Dans ce cas, la *relation ressentie* de l'agent va tendre vers la *relation idéale*. En revanche, l'interaction est dite *non-congruente* si l'agent perçoit de la dominance chez l'utilisateur. L'agent peut alors adopter trois différentes stratégies : dans le cas (a) il va changer sa *relation ressentie* envers l'utilisateur, dans le cas (b) il va changer sa propre *relation idéale* et dans le cas (c) il va essayer de changer l'attitude exprimée par l'utilisateur. Finalement, cette recherche de congruence va dépendre des stratégies adoptées par l'agent.

5. Conclusion et perspectives

Dans cet article, nous avons proposé un modèle conceptuel des attitudes sociales d'un agent virtuel. L'*attitude sociale exprimée* est ainsi déterminée par des aspects spontanés et stratégiques, dépendant entre autres de la personnalité et du rôle de l'agent. L'intercompréhension entre l'humain et la machine n'étant pas suffisante pour garantir la création et le maintien d'une relation à long terme, nous nous appuyons sur la théorie de la complémentarité. La dynamique de ces relations est ainsi liée à la recherche de cette complémentarité, modélisée par la théorie de la congruence : si l'attitude perçue par l'agent n'est pas conforme à celle qu'il désire percevoir, ce dernier va devoir adopter une stratégie en conséquence.

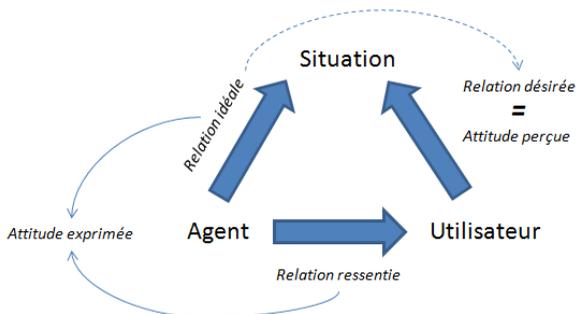


Figure 1 : Exemple d'interaction congruente

La modélisation de ces stratégies fait partie de nos travaux en cours. Nous avons en effet besoin de savoir comment les relations sociales influencent la prise de décision d'un agent. Dans (Raven, 2008), l'auteur dresse une liste de stratégies dominantes basée sur les coûts et récompenses estimés de chaque action, tout en prenant en compte les buts de l'individu. Dans (Schulman & Bickmore, 2009), les auteurs constatent qu'il est plus aisé d'être influencé par quelqu'un que l'on apprécie, si tant est que la décision à prendre n'ait qu'une importance mineure (Bickmore & Picard, 2005). Il peut ainsi être considéré plus pertinent pour l'agent de se lier d'amitié avec l'utilisateur s'il veut essayer de l'influencer plus efficacement.

Remerciements

Les travaux de recherche présentés dans cet article sont financés par le projet ANR MoCA.

Bibliographie

Ballin, D., Gillies, M., & Crabtree, I. (2004). A Framework For Interpersonal Attitude And Non-Verbal Communication In Improvisational Visual Media Production. *First European Conference on Visual Media Production IEE*. London, UK.

Baron-Cohen, S., Leslie, A., & Frith, U. (1985, Octobre). Does the Autistic Child Have a "Theory of Mind" ? *Cognition*, vol. 21, n°1 , pp. 37-46.

Bickmore, T., & Picard, R. (2005). Establishing and maintaining long-term human-computer relationships. *ACM Transactions on Computer-Human Interaction (TOCHI)* vol.12, n°2 , 293-327.

Bledsoe, W. (1986). I Had a Dream: AAI Presidential Address, 19 August 1985. *AI Magazine* vol.7, n°1 , 57-61.

Cafaro, A., Vilhjálmsón, H., Bickmore, T., Heylen, D., Jóhannsdóttir, K., & Valgarðsson, G. (2012). First Impressions: Users' Judgments of Virtual Agents' Personality and Interpersonal Attitude in First Encounters. *IVA'12*. Santa Cruz, UA.

Cassell, J. (2000). More Than Just Another Pretty Face: Embodied Conversational Interface Agents. *Communications of the ACM* vol.43, n°4 , 70-78.

Gatica-Perez, D. (2009). Automatic nonverbal analysis of social interaction in small groups: A review. *Image and Vision Computing* vol.27, n°12 , 1775-1787.

Grandgeorge, M., & MIAC. (2011). L'intercompréhension dans les interactions: modélisation interdisciplinaire autour de l'homme, l'animal et la machine. *III 2011*. Vannes, France.

Gurtman, M. (2009). Exploring Personality with the Interpersonal Circumplex. *Social and Personality Psychology Compass* vol.3, n°4 , 601-619.

Krämer, N. (2008). Social effects of virtual assistants. A review of empirical results with regard to communication. *Proceedings of the 8th international conference on Intelligent Virtual Agents (IVA'08)* (pp. 507-508). Springer-Verlag.

Marsella, S., Pynadath, D., & Read, S. (2004). PsychSim: Agent-based modeling of social interactions and influence. *Proceedings of the International Conference on Cognitive Modeling*.

Moshkina, L., & Arkin, R. (2003). On TAMEing robots. *IEEE International Conference on Systems, Man and Cybernetics*, 4, 3949-3959.

Ochs, M., Bevacqua, E., Prepin, K., Le, Q., Ding, Y., Huang, J., et al. (2011). La compréhension machine à travers l'expression non-verbale. Dans B. L.-S. Marine Grandgeorge, *Intercompréhensions comparées dans trois types d'interactions : Homme-Homme, Animal-Homme-Machine et Homme-Machine*. E.M.E Editions.

Osgood, C., & Tannenbaum, P. (1955). The Principle of Congruity in the Prediction of Attitude Change. *Psychological Review* vol.62, n°1, 42-55.

Prada, R., & Paiva, A. (2008). Social Intelligence in Virtual Groups. *Studies in Computational Intelligence* vol.140, 113-132.

Raven, B. (2008). The Bases of Power and the Power/Interaction Model of Interpersonal Influence. *Analyses of Social Issues and Public Policy* vol.8, n°1, 1-22.

Reeves, B., & Nass, C. (1996). *The Media Equation: How People Treat Computers, Television, and New Media Like Real People and Places*. Cambridge University Press.

Rist, T., Schmitt, M., Pelachaud, C., & Bilvi, M. (2003). Towards a Simulation of Conversation with Expressive Embodied Speakers and Listeners. *Proceedings of the 16th International Conference on Computer Animation and Social Agents (CASA'03)*.

Sadler, P., Ethier, N., & Woody, E. (2011). Interpersonal Complementarity. Dans *Handbook of interpersonal psychology: Theory, research, assessment, and therapeutic interventions* (pp. 123-142). New York: Horowitz, L.M. ; Strack, S.

Scherer, K. (2005). What are emotions? And how can they be measured? *Social Science Information* vol.44, n°4, 695-729.

Schulman, D., & Bickmore, T. (2009). Persuading users through counseling dialogue with a conversational agent. *PERSUASIVE'09*. Claremont, USA.

Swartout, W., & al, &. (2010). Ada and Grace: Toward Realistic and Engaging Virtual Museum Guides. *Proceedings of the 10th international conference on Intelligent virtual agents (IVA'10)*, (pp. 286-300).

Tracey, T. (2005). Interpersonal rigidity and complementarity. *Journal of Research in Personality* vol.39, n°6, 592-614.

Trapnell, P., & Wiggins, J. (1990). Extension of the Interpersonal Adjective Scales to Include the Big Five Dimensions of Personality. *Journal of Personality and Social Psychology* vol.59, n°4, 781-790.

Wiggins, J., Trapnell, P., & Phillips, N. (1988). Psychometric and geometric characteristics of the Revised Interpersonal Adjective Scales (IAS-R). *Multivariate Behavioral Research* vol.23, n°4, 517-530.