
Sujet de thèse - campagne 2020

Laboratoire : Institut des Systèmes Intelligents et de Robotique - ISIR

Etablissement de rattachement : Sorbonne Université

Titre de la thèse : **Amélioration de l'interaction humain/agent virtuel par le développement d'une boucle interactive humain/agent à partir de réseaux de neurones récurrents**

Directeur de thèse : Catherine Achard

Mail de contact : catherine.achard@sorbonne-universite.fr

Codirection éventuelle : Catherine Pelachaud

Collaborations dans le cadre de la thèse :

Rattachement à un programme :

Cotutelle envisagée :

Si oui avec quelle université & quel laboratoire :



Le sujet peut-il être publié sur le site web de l'ED SMAER : OUI

Résumé du sujet :

L'objectif de cette thèse est de développer la modélisation de la bouche interactive agent/humain à partir de réseaux de neurones récurrents. L'idée est d'inclure dans l'architecture une analyse et une modélisation prenant en compte de nombreux signaux sociaux comme le sourire, les mouvements de tête, l'expressivité corporelle ou la posture avec un accent mis sur la modélisation de l'imitation des signaux sociaux (posture, sourire...). Plusieurs architectures de réseaux pourront être testées comme par exemple un simple LSTM tenant compte des signaux sociaux des deux partenaires ou des LSTM couplés (un par partenaires) interagissant entre eux.

Les résultats seront intégrés sur la plateforme GRETA mettant en œuvre un agent virtuel à travers un écran géant.

Amélioration de l'interaction humain/agent virtuel par le développement d'une boucle interactive humain/agent à partir de réseaux de neurones récurrents

	Sorbonne Université Faculté des Sciences et d'Ingénierie
	Institut des Systèmes Intelligents et de Robotique UMR 7222 (ISIR) T55-65 Pyramide 4 place Jussieu, 75005 PARIS
Contacts	Catherine Achard catherine.achard@sorbonne-universite.fr Catherine Pelachaud catherine.pelachaud@upmc.fr

Contexte: Dans l'interaction humaine, l'homme s'adapte et ajuste son comportement en fonction de celui de ses interlocuteurs¹. Par exemple, un auditeur acquiesce pour indiquer son accord avec l'interlocuteur, il regarde le même objet ou sourit en réponse au sourire de son interlocuteur. Certains comportements comme le sourire mais aussi le niveau de synchronisation et d'imitation sont souvent des marqueurs de la qualité de l'interaction et du niveau d'engagement des participants de l'interaction. Dans cette optique, nous cherchons à modéliser un agent virtuel conversationnel animé capable d'adapter son comportement en fonction de celui de l'utilisateur.



Les comportements non verbaux jouent un rôle important pour maintenir l'engagement entre les humains et l'agent^{2,3}. C'est pourquoi nous sommes particulièrement intéressés par l'adaptation dynamique des comportements non verbaux de l'agent face à ceux de son interlocuteur et développons, depuis plusieurs années, une plateforme d'agent virtuel GRETA⁴, dotée d'un logiciel d'analyse multimodale du comportement de l'utilisateur EyesWeb⁵, d'un gestionnaire de dialogue Flipper⁶ et d'un module de génération de comportements⁴.

Pour mieux adapter le comportement de l'agent à celui de l'utilisateur, nous souhaitons utiliser les avancées récentes sur l'apprentissage et plus particulièrement celles du domaine des réseaux de neurones. Une étude récente que nous avons menée⁷, utilisant les réseaux de neurones récurrents,

¹ Judee K. Burgoon, Lesa A. Stern, and Leesa Dillman. 2010. Adaptation in Dyadic Interaction: Defining and Operationalizing Patterns of Reciprocity and Compensation. *Communication Theory* 1993 (2010).

² Timothy Bickmore, Daniel Schulman, and Langxuan Yin. 2012. Maintaining Engagement in Long-term Interventions with Relational Agents. *International Society of Differentiation* 83, 2 (2012), 1–29

³ Huang, L., Morency, L. P., & Gratch, J. (2011, September). Virtual Rapport 2.0. In *International workshop on intelligent virtual agents* (pp. 68-79). Springer, Berlin, Heidelberg.

⁴ Florian Pecune, Angelo Cafaro, Mathieu Chollet, Pierre Philippe, and Catherine Pelachaud. 2014. Suggestions for Extending SAIBA with the VIB Platform. In *Workshop on Architectures and Standards for IVAs, held at the '14th International Conference on Intelligent Virtual Agents (IVA 2014)*. Bielefeld eCollections, Boston, MA, USA, 16–20.

⁵ G. Volpe, P. Alborno, A. Camurri, P. Coletta, S. Ghisio, M. Mancini, R. Niewiadomski, and S. Piana. 2016. Designing Multimodal Interactive Systems Using EyesWeb XMI. *SERVE@AVI* (2016), 49–56.

⁶ Jelte van Waterschoot, Merijn Bruijnes, Jan Flokstra, Dennis Reidsma, Daniel Davison, Mariët Theune, and Dirk Heylen. 2018. Flipper 2.0. In *Proceedings of the 18th International Conference on Intelligent Virtual Agents*. 43–50.

⁷ Dermouche, Soumia and Catherine Pelachaud. "Generative Model of Agent's Behaviors in Human-Agent Interaction." *ICMI '19* (2019).

Sciences Mécaniques, Acoustique, Electronique, Robotique

montre que la modélisation temporelle conjointe des comportements non verbaux de l'agent et de l'utilisateur améliore l'interaction, grâce à une adaptation permanente du comportement de l'agent. Cette adaptation, qui crée une boucle interactive entre l'humain et l'agent, a été validée grâce à des questionnaires et nous aimerions l'améliorer afin de prendre en compte toutes les comportements de l'agent et optimiser la qualité de l'interaction.

Objectif de la thèse :

L'objectif de cette thèse est de continuer à développer la modélisation de la bouche interactive agent/humain à partir de réseaux de neurones récurrents. Peu de travaux, à notre connaissance, fondés sur les réseaux de neurones, permettent aujourd'hui de générer une interaction temps-réelle continue. Gordon et al.⁸ utilisent un algorithme de renforcement pour générer des expressions faciales et Dermouche et al.⁷ emploient les LSTM pour générer le sourire, la direction du regard et les mouvements de tête de l'agent en fonction des comportements de l'agent et de ceux de son interlocuteur humain.

Nous souhaitons étendre ces derniers travaux pour contrôler un plus grand nombre de comportements non verbaux de l'agent grâce à une boucle interactive, gage, nous l'espérons, d'une interaction de qualité. Celle-ci sera mesurée objectivement (niveau de synchronisation des interlocuteurs⁹ et d'imitation) et subjectivement (par questionnaires). L'idée est d'inclure une analyse et une modélisation plus fine en prenant en compte dans la modélisation plus de signaux sociaux comme le sourire, les mouvements de tête, l'expressivité corporelle et la posture avec un accent mis sur la modélisation de l'imitation de ces signaux sociaux. Plusieurs architectures de réseaux pourront être testées et évaluées comme par exemple les LSTM ou les réseaux convolutionnels temporels de manière à prendre en compte la temporalité des signaux, leur synchronie, ...

Les résultats escomptés visent à être intégrés sur la plateforme GRETA mettant en œuvre un agent virtuel à travers un écran géant. Le but est de fournir une démonstration de l'amélioration de cette plateforme au travers d'une application telle qu'un agent d'accueil ou un agent guide.

Cadre applicatif :

Cette thèse vise à développer un agent virtuel, doté de compétences sociales, en vue d'améliorer la qualité des interactions. Il pourra être utilisé dans le cadre d'un assistant virtuel ou être intégré dans des jeux sérieux.

Les interactions humain/agent ainsi générées seront évaluées grâce à des questionnaires sur la perception de l'agent et de l'interaction. On considèrera aussi un questionnaire¹⁰ sur les *a priori* des participants face aux nouvelles technologies et en particulier, face aux agents sociaux.

Techniques utilisées : Apprentissage, deep learning

Qualités du candidat requises : Titulaire d'un M2 d'Ingénierie ou d'informatique ou d'un diplôme d'ingénieur, possédant des compétences en apprentissage et souhaitant explorer un sujet théorique à l'interface de l'informatique et des sciences comportementales

⁸ Goren Gordon, Samuel Spaulding, Kory Westlund, Joo Lee, Luke Plummer, Marayna Martinez, and Madhurima Das. 2016. Affective Personalization of a Social Robot Tutor for Children 's Second Language Skills. 2011 (2016), 3951–3957.

⁹ Delaherche, E., Chetouani, M., Mahdhaoui, A., Saint-Georges, C., Viaux, S., & Cohen, D. (2012). Interpersonal synchrony: A survey of evaluation methods across disciplines. *IEEE Transactions on Affective Computing*, 3(3), 349-365.

¹⁰ T. Nomura, T. Kanda, and T. Suzuki. 2006. Experimental Investigation into Influence of Negative Attitudes toward Robots on Human-Robot Interaction. *AI & Society* 20, 2 (2006).