

## Version française

# Modélisation de l'interaction social en action conjointe humain-robot à partir d'une approche bio-inspirée

**Directrice de thèse** : Laure Buhry (MCF-HDR – LORIA)

**Co-encadrant** : Hendry F. Chame (MCF – LORIA)

**Mots-clés.** Interaction Humain-robot, Action conjointe, Attention conjointe, Modélisation bio-inspirée, Robotique neuronale, Cognition sociale.

## Contexte et état de l'art

La robotique sociale est une aire de recherche pluridisciplinaire intéressée par l'étude de l'interaction des robots au niveau social. Ces dernières décennies, la recherche sur les interactions homme-robot (IHR) s'est accompagnée de la conception de robots de plus en plus habiles pour s'engager dans des actions conjointes avec l'humain. Cependant, ces progrès ont rencontré aussi des défis importants pour assurer des interactions fluides et maintenir l'engagement de l'humain (Belhassein et al. (2022)).

A partir de la perspective de recherche en cognition 4E (Newen et al. (2018)), notamment inspirée de la théorie de l'interaction (Gallagher (2008)), un aspect auquel ce projet s'intéresse est la modélisation de situations d'interaction directes et intuitives humain-robot. Notamment, il s'agit de situations caractérisées par un échange basé sur un ensemble de compétences cognitives primaires (e.g le suivi du regard, l'imitation, le partage attentionnel), dont la communication verbale bien que possible reste moins prioritaire.

En particulier, ce projet s'intéresse à l'étude de l'*action conjointe* en IHR, considéré selon Fiebich et Gallagher (2013) comme une forme complexe d'interaction sociale, caractérisée par les trois éléments suivants : a) une intention partagée (c'est-à-dire viser un objectif commun), b) la connaissance commune de viser le même objectif, et c) la participation à des patrons de comportement coopératifs, où les aspects motivationnels et les processus attentionnels conjoints revêtent une extrême importance.

Dans le domaine de l'interaction homme-machine, Vesper et al. (2010) ont proposé une description architecturale minimale et modulaire des types de processus dédiés permettant l'action conjointe. Ce projet de recherche s'intéresse à l'intégration entre la modélisation architecturale et l'étude de l'interaction en tant que système dynamique, capable de représenter de façon adéquate l'aspect rythmique (e.g. la synchronisation) du comportement et de l'activité cérébrale de l'humain en

correspondance à l'état interne et actions du robot. Cette possibilité d'intégration sera poursuivie à partir de la perspective de la théorie du *principe de l'énergie libre*, notamment le concept d'*inférence active* (Friston et al. (2013), Allen et Friston (2018)).

## **Pertinence, originalité et objectifs**

Ce travail a pour objectif l'étude des aspects dynamiques de l'action conjointe en interaction humain-robot, grâce à des approches de modélisation mathématique, de simulation et de traitement du signal multi-échelles, et de prototypage des expériences d'interaction. Par conséquent, l'objectif principal de ce projet de recherche est la proposition d'un modèle mathématique permettant de représenter et suivre en temps-réel la dynamique de l'interaction. La modélisation résultante pourrait contribuer d'une part au développement de méthodes diagnostiques destinées à estimer la qualité de l'interaction en IHR. D'une autre part, vue le niveau d'activité d'interaction envisagé (un échange intuitif entre l'humain et le robot), ce projet pourrait contribuer au développement de méthodes destinés à étudier la cognition sociale en plusieurs conditions psychologiques (e.g. dans le cadre du trouble du spectre de l'autisme, la condition de réhabilitation cognitive, ou des conditions physiopathologiques comme la schizophrénie).

## **Méthodologie et techniques mises en œuvre**

*Définition de l'activité d'interaction.* En se basant sur la littérature, la première étape consistera à proposer une situation d'interaction intuitive humain-robot, incluant une activité susceptible d'inciter l'action conjointe entre les partenaires. Idéalement, cette activité se présentera sous la forme de jeu amusant, ayant le potentiel de favoriser l'engagement de l'humain.

*Modélisation mathématique bio-inspirée.* La seconde étape visera à la construction du modèle d'interaction en temps-réel à partir de la définition du modèle de données. Le travail s'appuiera sur une approche à la modélisation de systèmes dynamiques inspirée du vivant (e.g. les modèles de réseaux neuronaux d'attracteurs continus inspirés de la *théorie des champs neuronaux dynamiques*, Amari (1977)). En suivant cette approche, des travaux précédents ont considéré deux aspects directement reliés à la problématique en étude, tels que la modélisation de l'*attention conjointe* de plusieurs individus en interaction (Chame et al. (2023)) et la modélisation de l'état motivationnel en interaction humain-système selon la théorie de l'autodétermination (Chame et al. (2019)). D'autres travaux qui pourraient servir de référence à la modélisation de l'interaction se sont inspirés du concept d'inférence active issue de la théorie du principe de l'énergie libre (e.g. Chame et Tani (2020), Chame et al. (2020)).

*Prototypage de l'expérience.* Afin de fournir les données au modèle d'interaction, le projet s'appuiera sur la prise de données électrophysiologiques (e.g. électroencéphalogramme) et comportementales (e.g. suivie de la position, la direction du regard), ainsi que les données capturées du robot (son état interne provenant des capteurs proprioceptives et extéroceptives). En raison des contraintes technologiques, le prototype d'interaction consistera probablement à un système distribué, dont la capture de données doit être synchronisée sur horloge. Ainsi, le système

sera implémenté dans le cadre intergiciel Robot Operating System (ROS), programmé en langage Python ou C++.

### **Encadrement et collaborations**

La thèse se déroulera au LORIA (Neurorhythms). Il est prévu une collaboration avec Rachid Alami, membre du LAAS-CNRS (RobotS and InteractionS - RIS), pour ce qui concerne la mise en œuvre de l'expérience réelle. Les expertises du LORIA seront particulièrement sollicitées pour les aspects de modélisation et de simulation, en particulier pour la modélisation de l'interaction biologiquement inspirée. Alors que l'expertise du LAAS-CNRS sera sollicitée dans la phase expérimentale.

## English Version

# Modeling social interaction in human-robot joint action from a bio-inspired approach

**Thesis director** : Laure Buhry (MCF-HDR – LORIA)

**Co-supervision** : Hendry F. Chame (MCF – LORIA)

**Keywords.** Human-robot interaction, Joint Action, Joint Attention, Bio-inspired modeling, Neural robotics, Social cognition.

## Context and state of the art

Social robotics is a multidisciplinary research field interested in studying the interaction of robots at the social level. In recent decades, research on human-robot interactions (HRI) has been accompanied by the design of robots that are increasingly more capable at engaging in joint actions with humans. However, these advances have also encountered significant challenges in ensuring smooth interactions and maintaining human engagement (Belhassein et al. (2022)).

From the perspective of research in 4E cognition (Newen et al. (2018)). Notably, inspired by *interaction theory* (Gallagher (2008)), one aspect in which this project is interested is the modeling of direct and intuitive HRI situations. These are for instance situations characterized by an exchange based on a set of primary cognitive skills (e.g. gaze tracking, imitation, attentional sharing), of which verbal communication, although possible, remains less of a priority.

In particular, this project is interested in the study of Joint Action in IHR, considered according to Fiebich and Gallagher (2013) as a complex form of social interaction, characterized by the following three elements: a) a shared intention (i.e. aiming for a common goal), b) common knowledge of aiming for the same goal, and c) participation in cooperative patterns of behavior, where motivational aspects and joint attentional processes are of extreme importance.

In the field of HRI, Vesper et al. (2010) proposed a minimal and modular architectural description of the types of dedicated processes enabling Joint Action. This research project focuses on the integration between architectural modeling and the study of interaction as a dynamical system, capable of adequately representing the rhythmic aspect (e.g. synchronization) of behavior and brain activity of the human in correspondence to the internal state and actions of the robot. This possibility of integration will be pursued from the perspective of *free energy principle* theory through the concept of *active inference* (Friston et al. (2013), Allen and Friston (2018)).

## Relevance, originality and objectives

This research project aims at studying the dynamics of Joint Action (JA) in human-robot interaction (HRI) through mathematical modeling, simulations, multi-scale signal processing and prototyping

human-robot interaction experiments. Therefore, the main objective of the project is the proposal of a mathematical model allowing to represent and track in real time JA. The resulting model could contribute, on the one hand, to the development of diagnostic methods intended to estimate the quality of interaction in HRI. On the other hand, given the level of interaction activity envisaged (i.e. intuitive interaction between the human and the robot), this project could contribute to the development of methods intended to study social cognition in several psychological conditions (e.g. autism spectrum disorder, cognitive rehabilitation condition, or psychopathological conditions such as schizophrenia).

## **Methodology**

*Definition of the interaction activity.* Based on the literature, the first step will consist in proposing a situation of intuitive human-robot interaction, including an activity having the potential to incite Joint Action between the partners. Ideally, this activity will take the form of a fun game, capable of inducing human engagement.

*Bio-inspired mathematical modeling.* The second step will aim at the construction of the real-time interaction model from the definition of the data model. The model proposition will be bio-inspired (e.g. *continuous attractor neural network models* are inspired by the theory of dynamic neural fields, Amari (1977)). Following this approach, previous work has considered two aspects directly related to the problem under study, such as the modeling of Joint Attention for several interacting individuals (Chame et al. (2023)) and the modeling of the motivational state in human-system interaction, according to self-determination theory (Chame et al. (2019)). Other works could inspire the modeling phase, such as the concept of *active inference* from free energy principle theory (e.g. Chame and Tani (2020), Chame et al. (2020)).

*Experiment prototyping.* In order to provide data to the interaction model, the project will rely on tracking electrophysiological (e.g. electroencephalogram) and behavioral (e.g. monitoring of position, direction of gaze) data, as well as data captured from the robot (its internal state from proprioceptive and exteroceptive sensors). Due to technological constraints, the interaction prototype will probably consist of a distributed system, whose data capture must be clock synchronized. Thus, the system will be implemented in the Robot Operating System (ROS) middleware, programmed in Python or C++.

## **Supervision and collaborations**

The thesis will take place at LORIA (Neurorhythms team). A collaboration with Dr Rachid Alami, member of LAAS-CNRS (RobotS and InteractionS - RIS) is expected for the experimental phase. The expertise of LORIA will be particularly required for the modeling and simulation aspects, in particular for the modeling of the biologically inspired interaction. While the expertise of LAAS-CNRS will be requested in the experimental phase.

## References

- Amari, S. I. (1977). Dynamics of pattern formation in lateral-inhibition type neural fields. *Biological cybernetics*, 27(2), 77-87.
- Allen, M., & Friston, K. J. (2018). From cognitivism to autopoiesis: towards a computational framework for the embodied mind. *Synthese*, 195(6), 2459-2482.
- Belhassein, K., Fernández-Castro, V., Mayima, A., Clodic, A., Pacherie, E., Guidetti, M., Alami, R., and Cochet, H. (2022). Addressing joint action challenges in HRI: Insights from psychology and philosophy. *Acta Psychologica*, 222, 103476.
- Chame, H. F., Mota, F. P., & da Costa Botelho, S. S. (2019). A dynamic computational model of motivation based on self-determination theory and CANN. *Information Sciences*, 476, 319-336.
- Chame, H. F., & Tani, J. (2020, May). Cognitive and motor compliance in intentional human-robot interaction. In *2020 IEEE International Conference on Robotics and Automation (ICRA)* (pp. 11291-11297). IEEE.
- Chame, H. F., Ahmadi, A., & Tani, J. (2020). A hybrid human-neurorobotics approach to primary intersubjectivity via active inference. *Frontiers in Psychology*, 11, 584869.
- Chame, H. F., Clodic, A., & Alami, R. (2023, May). TOP-JAM: A bio-inspired topology-based model of joint attention for human-robot interaction. In *2023 IEEE International Conference on Robotics and Automation (ICRA)*.
- Fiebich, A., & Gallagher, S. (2013). Joint attention in joint action. *Philosophical Psychology*, 26(4), 571-587.
- Friston, K., Schwartenbeck, P., FitzGerald, T., Moutoussis, M., Behrens, T., & Dolan, R. J. (2013). The anatomy of choice: active inference and agency. *Frontiers in human neuroscience*, 7, 598.
- Gallagher, S. (2008). "Understanding others: embodied social cognition," in *Handbook of Cognitive Science*, eds P. Calvo and A. Gomila (San Diego, CA; Oxford; Amsterdam: Elsevier), 437–452.
- Newen, A., Bruin, L., and Gallagher, S. (2018). "4E cognition: historical roots, key concepts and central issues," in *The Oxford Handbook of 4E Cognition*, eds A. Newen, S. Gallagher, and L. de Bruin (Oxford: Oxford University Press), 3–15.
- Vesper, C., Butterfill, S., Knoblich, G., & Sebanz, N. (2010). A minimal architecture for joint action. *Neural Networks*, 23(8-9), 998-1003.