

Analyse de données complexes et hétérogènes pour l'adaptation de scénarios robotiques avec la prise en compte de toutes les dimensions de l'interaction

Contexte du projet :

A travers cette thèse, nous souhaitons répondre à une problématique sociétale soulevée dès 2016 à travers une étude du ministère de la culture qui indique que le jeune public déserte les lieux de culture avec en particulier les musées.

La transformation numérique des musées a été amorcée ces dernières années pour répondre à ces problématiques. A travers cette thèse, nous souhaitons accompagner cette transformation, ce qui s'intègre pleinement dans un des quatre domaines d'expertise de l'Université (transition numérique).

Outre l'aspect culturel, nous souhaitons également attirer l'attention du public sur les questions de l'environnement à travers des jeux sérieux sur l'ensemble de la ville de La Rochelle. L'objectif ici est de déployer plusieurs dispositifs (robots, bornes, etc.) à l'échelle de la ville via lesquels le public se sensibilisera sur ces questions.

Cette thèse s'inscrit également dans une volonté de développer des outils d'analyse interactive des données qui mettraient l'analyste de données au cœur du processus.

Description du sujet :

De nombreuses approches de conception d'applications robotiques ont été proposées ces dernières années [Dhouib2012, Miyazawa2016, Miyazawa2019] et si elles s'avèrent pertinentes pour le problème qu'elles résolvent, la plupart sont dépendantes de la plateforme ou du langage utilisé. Pour résoudre ce problème, l'utilisation de modèles formels pour la génération de scénarios ou d'activités est essentielle pour garantir une expérience de haute qualité. Dans ce contexte, le modèle formel CIT (Content Interaction Time) [Mondou2018, Mondou2019], consacré au développement d'applications robotiques interactives et basé sur des réseaux d'automates temporisés à entrées/sorties a été proposé. Ce modèle est capable de prendre en compte les dimensions de temps, de contenus et d'interactions en facilitant la phase de conception d'un scénario interactif et répondant aux critères de généricité, de modularité, d'extensibilité et de temporalité.

Les premiers travaux cités précédemment ont permis de mener différentes expérimentations dans deux musées partenaires (Muséum d'Histoire Naturelle de La Rochelle et Musée Sainte-Croix de Poitiers). Celles-ci nous ont montré, à travers la diversité du public et de son appréhension de l'interaction avec le robot, que le paramétrage des interactions conditionne leur bonne réalisation. Plus précisément, nous avons identifié deux verrous scientifiques pour améliorer ce paramétrage et qui sont l'objet de cette thèse, à savoir (1) l'intégration de l'espace dans le modèle, et (2) une adaptativité dynamique du modèle.

- **Intégration de l'espace dans le modèle**

Si ce modèle à deux couches permet de concevoir une activité interactive de façon modulaire, les différentes interactions ne sont pas liées à un espace particulier dans lequel elles doivent se dérouler. Les robots et les joueurs étant mobiles, l'intégration de l'espace dans le modèle représente une perspective d'amélioration majeure. L'objectif ici est donc de pouvoir spécifier le lieu dans lequel l'action doit se dérouler. Cette future gestion lève la problématique de la cohérence et la représentation de l'information de localisation dans le modèle. Doit-on l'externaliser ou l'intégrer à la définition du comportement, de l'agent ou du contexte ?

Cette notion de gestion de l'espace pourra être intégrée au modèle grâce aux travaux de suivi des déplacements de personnes en intérieur qui est traité au L3i depuis quelques années et plus particulièrement dans une thèse en cours. L'objectif est pour cette thèse d'être capable de remonter de l'information en continu au serveur sur lequel le robot va puiser les données pour décider de telle ou telle action. L'approche développée au L3i consiste à la mise en place d'un système de micro-localisation intérieure à base de balises BLE. Il a été pensé pour être facilement et rapidement déployable dans la plupart des environnements fermés. Les algorithmes proposés sont capables de reconstruire en temps quasi réel la séquence de déplacements d'une personne [Richard2021]. La thèse s'appuiera sur ces travaux et à l'intégration de ces résultats afin d'ajouter la notion d'espace au modèle CIT dans un premier temps pour une gestion en milieu fermé et dans un deuxième temps, à l'échelle d'un espace ouvert (territoire d'une ville).

- **Adaptativité dynamique du modèle**

Nous avons obtenu, à travers les différentes expériences publiques passées, un ensemble de traces d'exécution. Nous souhaitons donc, à travers cette thèse, proposer une adaptation dynamique du scénario en fonction du comportement réel de l'utilisateur (adaptation des paramètres de temps, de la vitesse de parole du robot, etc.).

L'adaptation des comportements d'un système lors de son exécution est une tâche difficile et largement développée ces dernières années dans l'état de l'art [Gonçalo2019, Krupitzer2015]. La difficulté principale est due à l'absence de méthodes précises pour la description et la modélisation des systèmes. Dans le domaine de la robotique, de nombreuses approches utilisent l'apprentissage par renforcement pour répondre à cette problématique. Si l'intégration d'un algorithme d'apprentissage par renforcement a eu lieu dans le modèle CIT, celui-ci n'a pas encore permis de répondre à nos attentes. Une telle approche se positionne parmi les méthodes dites « boîtes noires » qui fournissent souvent de très bons résultats mais avec un pouvoir d'explicabilité limité, ce qui rend plus difficilement exploitables les résultats obtenus. À l'inverse, les méthodes dites « boîtes blanches » offrent explicabilité et lisibilité, une lisibilité qui peut s'adresser à un expert pour une adaptation du modèle.

Des résultats récents et novateurs obtenus au laboratoire L3I permettent d'envisager de proposer de nouvelles méthodes d'analyse de données complexes et hétérogènes offrant à la fois lisibilité et interactivité. Il s'agit d'un nouvel algorithme, NextPriorityConcept [Demko20] qui propose de générer des regroupements de données appelés « concepts » à partir de données complexes et hétérogènes, et de la plateforme GALACTIC qui en fournit une implémentation générique pour de nouvelles méthodes d'analyses de données interactives et explicatives, la première version ayant été diffusée en Janvier 2020. L'algorithme NextPriorityConcept se positionne dans le domaine scientifique de la fouille de motifs et de l'Analyse de Concepts Formels [Ganter99] qui met en avant la structure de « treillis des concepts » où un concept est une paire composée d'un sous-groupe maximal d'objets, avec les données qu'ils partagent. L'ensemble de tous les concepts s'organise naturellement sous la forme d'un graphe hiérarchique, appelé treillis des concepts, offrant ainsi un support pour visualiser et interagir avec une représentation des données. NextPriorityConcept est une extension de la génération du treillis des concepts pour des données non binaires, complexes et hétérogènes, où chaque concept est composé d'un sous-groupe de données avec descriptions commune.

GALACTIC (GAlois LAttices, Concept Theory, Implicational systems and Closures) est une plateforme de développement qui propose une implémentation générique de l'algorithme NextPriorityConcept, et un écosystème d'extensions pour définir des types de données, leurs descriptions, et des stratégies d'exploration. Des extensions sont implémentées pour des données binaires, catégorielles, numériques, des chaînes de caractères ainsi que des séquences temporelles avec un indice temporel ou encore avec un intervalle de temps. Les extensions pour des séquences ont été développées dans le cadre d'un projet d'analyse de déplacements de touristes et de visiteurs de musées [Boukhetta21]. Plusieurs séquences ont pu être extraites à partir des trajectoires des touristes (séquences géographique, séquences enrichies avec des données de contexte telles que la météo, la marée), une analyse multi-séquences avec Galactic a permis de faire émerger des comportements pertinents. Un projet avec la SATT Aquitaine est prévu en 2022 pour une montée en maturation de GALACTIC. L'objectif est de développer une interface logicielle interactive de navigation dans le treillis, permettant à l'analyste des données de visualiser les données de chaque concept en sélectionnant les descriptions voulues, mais aussi de les affiner en choisissant différentes stratégies d'exploration des données. Il s'agit d'une navigation dans les concepts du treillis pour une recherche d'information interactive. À partir du concept initial contenant toutes les données, l'utilisateur

peut affiner sa recherche niveau par niveau jusqu'à atteindre le(s) concept(s) contenant les données qu'il recherche.

Ainsi, un concept pourrait ainsi correspondre à un sous-groupe de joueurs, décrit par leur comportement commun. Ce comportement commun pourrait intégrer des données hétérogènes issues à la fois du modèle CIT (paramètres de temps, de la vitesse de parole...), des trajectoires de déplacements (sous-séquences de déplacements communes, avec information temporelle), ainsi que des données de contexte (âge, nationalité, ...). L'interface logicielle interactive permettrait ainsi de visualiser ces différents concepts, mais également de les affiner selon différentes stratégies d'exploration (prise en compte ou non de la temporalité, de la localité, ...) dans un objectif d'adaptabilité du modèle CIT.

Informations complémentaires :

Encadrants :

- Directeur de thèse : Karell BERTET (50% d'encadrement)
- Encadrant scientifique : Damien MONDOU (50% d'encadrement)

Thématiques :

- Ingénierie des connaissances
- Analyse et gestion de contenus
- Interactivité et dynamique des systèmes

Date de début du contrat : Octobre 2022

Durée du contrat : 3 ans

Références bibliographiques :

[Boukhetta21] SE Boukhetta, C Demko, K Bertet, J Richard, C Cayère. Temporal Sequence Mining Using FCA and GALACTIC. International Conference on Conceptual Structures, 2021

[Demko20] C. Demko, K. Bertet, C. Faucher, J-F. Viaud, S. O. Kuznetsov. Next Priority Concept: A new and generic algorithm computing concepts from complex and heterogeneous data. Theoretical Computer Science. 845: 1-20 (2020)

[Dhouib2012] S. Dhouib, S. Kchir, S. Stinckwich, T. Ziadi & M. Ziane, RobotML, a Domain-Specific Language to Design, Simulate and Deploy Robotic Applications, in Simulation, Modeling, and Programming for Autonomous Robots (Berlin, Heidelberg) (I. Noda, N. Ando, D. Brugali & J. J. Kuffner, eds.), Springer Berlin Heidelberg, 2012, p. 149-160.

[Ganter99] Ganter, B., Wille, R.: Formal Concept Analysis. Mathematical foundations. Springer Verlag, Berlin, 1999.

[Gonçalo2019] Gonçalo S. Martins, Luís Santos & Jorge Dias, User-adaptive interaction in social robots: A survey focusing on non-physical interaction, International Journal of Social Robotics, 11(1) :185–205, Jan 2019.

[Krupitzer2015] Christian Krupitzer, Felix Maximilian Roth, Sebastian VanSyckel, Gregor Schiele & Christian Becker, A survey on engineering approaches for self- adaptive systems, Pervasive and Mobile Computing, 17 :184 – 206, 2015. 10 years of Pervasive Computing' In Honor of Chatschik Bisdikian.

[Miyazawa2016] A. Miyazawa, P. Ribeiro, W. Li, A. Cavalcanti, J. Timmis & J. Woodcock, RoboChart: a state-machine notation for modelling and verification of mobile and autonomous robots, Tech. Rep, 2016

[Miyazawa2019] A. Miyazawa, P. Ribeiro, W. Li, A. Cavalcanti, J. Timmis & J. Woodcock, RoboChart: modelling and verification of the functional behaviour of robotic applications, Software & Systems Modeling, 2019

[Mondou2018] D. Mondou, A. Prigent & A. Revel, A Dynamic Scenario by Remote Supervision: A SeriousGame in the Museum with a Nao Robot, in Advances in Computer Entertainment Technology(Cham)(A. D. Cheok, M. Inami & T. Romão, éd.), Springer International Publishing, 2018, p. 103-116

[Mondou2019] Damien Mondou, Gestion adaptative des contenus numériques : proposition d'un framework générique par apprentissage et re-scénarisation dynamique, 2019.

[Richard2021] J. Richard, K. Bertet, C. Faucher. Ble Based Indoor Positioning System and Minimal Zone Searching Algorithm (MZS) Applied to Visitor Trajectories within a Museum, *Applied Sciences*, 2021.

Contact:

Email (précisez dans l'objet [Candidature these]):

- damien.mondou@univ-lr.fr
- karell.bertet@univ-lr.fr