



HAL
open science

Apprentissage de comportements à partir de données temporelles hétérogènes

Nida Meddouri, François Rioult, Bruno Crémilleux

► **To cite this version:**

Nida Meddouri, François Rioult, Bruno Crémilleux. Apprentissage de comportements à partir de données temporelles hétérogènes. Conférence francophone sur l'Extraction et la Gestion des Connaissances, Association EGC, Jan 2022, Blois, France. pp.77-79. hal-03738180

HAL Id: hal-03738180

<https://hal.science/hal-03738180v1>

Submitted on 25 Jul 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Apprentissage de comportements à partir de données temporelles hétérogènes

Nida Meddouri*, François Rioult*, Bruno Crémilleux*

*UNICAEN, ENSICAEN, CNRS - UMR GREYC, Normandie Univ – 14000 Caen France
prenom.nom@unicaen.fr

Récemment, l'analyse et l'apprentissage du comportement d'un ou plusieurs agents (humains ou artificiels) a fait l'objet de plusieurs travaux de recherche (Darty et al., 2014) (Sun et al., 2021). La modélisation d'un comportement nécessite de concevoir et mettre en place de nouvelles approches et méthodes de fouille des données. Cette modélisation dépendra des observations d'un comportement, et donc d'une trace de données en tenant compte de l'aspect temporel (Dix et al., 1993).

Dans le cadre d'un projet pour étudier le comportement des apprentis soignants, les observations seront acquises d'un simulateur de réalité virtuelle modélisant la chambre d'un patient à l'hôpital. Des données hétérogènes – numériques et catégorielles – correspondent aux actions d'un apprenant en fonction de l'état des objets de la chambre. L'analyse par des techniques d'apprentissage des traces d'un apprenant et la modélisation de son comportement sont deux défis majeurs dans le cadre de ce projet.

1 Analyse de comportements à partir des données complexes et hétérogènes

La scène mise en oeuvre par le simulateur est une chambre de patient dans un hôpital. Elle contient potentiellement 142 objets. Parmi eux, sept sont mal placés, par exemple une seringue se trouve par terre. L'apprenant devra découvrir, dans un intervalle de temps fixe (une dizaine de minutes), ces objets aux placements erronés. Le simulateur permet à un apprenant de se déplacer librement dans la chambre, sélectionner un objet à la fois, signaler une erreur et justifier sa décision interactivement parmi une liste de choix. En fonction de sa position dans la chambre et des objets dans son champ de vision, un apprenant effectuera une action qui sera annotée par le simulateur : l'apprenant peut ignorer un objet, ou le sélectionner, soit pour le signaler et justifier une erreur, soit le remettre sans le signaler. Les coordonnées et les actions de l'apprenant ainsi que l'état des objets apparaissant dans le viseur du casque seront enregistrées à chaque instant d'une simulation. Ces traces représentent une séquence de données complexes (spatiales et temporelles) et hétérogènes (quantitatives et qualitatives).

Un comportement est un ensemble de réactions d'un sujet objectivement observables (Bloch et al., 2002). Il est considéré comme un choix tactique, évolue selon la dynamique de l'environnement et dépend de l'état mental de l'apprenant. Dans le cadre de ce travail, l'apprenant effectue une action en fonction des objets présents dans la chambre virtuelle et selon ses propres

Apprentissage de comportements à partir de données temporelles hétérogènes

règles (une stratégie). Par exemple, un apprenant aura tendance à inspecter les objets qui se trouvent à sa droite avec une priorité pour les objets en hauteur, ensuite les objets au sol, au fur et à mesure qu'il avance dans la chambre. Un autre apprenant aura tendance à aller directement vers le patient pour inspecter sa fiche médicale, en jetant un coup d'oeil sur la totalité de objets de la chambre, avant de commencer à inspecter les objets au voisinage du patient. Dans tous les cas, une séquence temporelle est une séquence d'actions, ordonnées chronologiquement, effectuées par un apprenant durant un intervalle de temps fini.

L'objectif de notre travail est d'apprendre un modèle comportemental à partir de traces d'interactions avec les objets. De plus, nous souhaitons faire émerger un modèle explicable et interprétable, pour échanger avec l'expert de la formation des apprentis soignants. Pour cela, nous proposons une formalisation logique du comportement d'un apprenant.

2 Approche proposée

Nous avons décidé d'extraire des règles d'induction (Cohen, 1995), qui ont la particularité d'être expressives, facile à générer, à interpréter. Ces règles sont de la forme Si condition Alors conclusion. Plus précisément, nous souhaitons focaliser notre attention sur des règles dont la condition représente une conjonction d'états de la simulation et la conclusion est une action entreprise par l'apprenant. Donnons deux exemples : (i) Si une seringue usée se trouve à droite et sur la table Alors l'apprenant va la saisir et la signaler ; (ii) Si le temps écoulé est inférieur une minute et l'apprenant s'approche du patient Alors il va ignorer tout les objets de la chambre. Nous considérons que le comportement d'un apprenant est un ensemble de règles d'induction.

L'étape suivante consiste à analyser et comparer les comportements de plusieurs apprenants pour enrichir le retour d'expérience auprès de l'expert. Nous proposons donc de regrouper les comportements similaires. Pour cela, nous allons générer un ensemble de règles associé à chaque comportement. Dans un premier temps, nous allons comparer les modèles générés en utilisant une mesure statistique comme indice de similarité. Cette mesure nous permet de construire des groupes homogènes d'apprenants à partir de comportements similaires. Enfin, il sera utile de caractériser chaque groupe par un comportement spécifique.

Pour pallier le problème de l'hétérogénéité des données, nous prévoyons de discrétiser les attributs quantitatifs, de façon à faire émerger des séquences de item ou d'ensembles d'items. Il nous semble raisonnable d'appliquer ensuite des algorithmes d'extraction de motifs séquentiels (Fournier Viger et al., 2017) disponibles dans la communauté¹.

3 Conclusion

Dans ce travail, chaque comportement d'apprenant est représenté par un ensemble de règles d'induction. L'analyse et la comparaison de ces ensembles permettront de grouper les comportements similaires des apprenants selon une mesure de similarité. Un comportement typique est détectable à partir des règles en commun dans les ensembles similaires.

Remerciements Ce travail est financé par le projet INCA – Interactions Naturelles avec des Compagnons Artificiels (région Normandie).

1. Par exemple <https://www.philippe-fournier-viger.com/spmf/>

Références

- Bloch, H., R. Chemama, et E. Dépret (2002). *Grand dictionnaire de la psychologie*. Collectif Larousse.
- Cohen, W. W. (1995). Fast effective rule induction. In *In Proceedings of the Twelfth International Conference on Machine Learning*, pp. 115–123. Morgan Kaufmann.
- Darty, K., J. Saunier, et N. Sabouret (2014). Analyse des comportements agents par agregation aux comportements humains.
- Dix, A., J. Finlay, et R. Beale (1993). Analysis of user behaviour as time series. In *Proceedings of the Conference on People and Computers VII, HCI'92, USA*, pp. 429–444. Cambridge University Press.
- Fournier Viger, P., C.-W. Lin, U. Rage, Y. S. Koh, et R. Thomas (2017). A survey of sequential pattern mining. *Data Science and Pattern Recognition 1*, 54–77.
- Sun, D., T. Li, F. You, M. Hu, et Z. Li (2021). Prediction of learning behavior characters of MOOC's data based on time series analysis. *Journal of Physics : Conference Series 1994*(1), 012009.

Summary

L'apprentissage d'un comportement à partir des données temporelles a fait l'objet de plusieurs travaux de recherche récemment. Les approches et les techniques de la fouille des données temporelles, sont des outils qui permettent de modéliser un comportement. Dans le cadre d'une chambre de patient dans un hôpital, nous proposons de sauvegarder une trace des actions d'un apprenant soignant et l'état de chaque objets dans cette chambre à chaque instant. Cette trace permettra de modéliser le comportement d'un apprenant sous forme d'un ensemble de règles d'induction. L'analyse et la comparaison des ensembles de règles générées permettra de comparer et regrouper les comportements de plusieurs apprenants et de déduire un comportement typique parmi ceux les plus similaires.