Sujet de stage M2
Février 2023

Comparaison de différents formalismes de causalité réelle : théorie, implémentations et illustrations

Informations pratiques

Lieu du stage : LIP6, Sorbonne Université, 4 place Jussieu, 75005 Paris
Encadrants : Isabelle Bloch, Gauvain Bourgne, Marie-Jeanne Lesot
Mail : prenom.nom@lip6.fr
Durée du stage : 6 mois

Mots clefs : Causalité, XAI, Ethique computationnelle, Langage d'action (PDDL),

Contexte

Dans le domaine de l'intelligence artificielle explicable (XAI), de nombreuses approches sont développées. Parmi celles-ci, une classe importante conçoit la recherche d'explications comme la recherche de causes à une observation ou une décision. Il s'agit donc d'aller au-delà de simples corrélations. Le but de ce stage est d'étudier les principaux formalismes dans cette classe de méthodes, en distinguant en particulier les méthodes qui recherchent toutes les causes possibles et celles qui cherchent à répondre à des questions particulières.

Travail de stage

Le travail se décompose en trois parties, fortement reliées entre elles, et qui seront donc menées en grande partie de manière parallèle :

1. analyse théorique de quelques approches clés, à partir des articles cités en référence ci-dessous, en mettant en évidence les principes des approches, leurs objectifs et leurs principales propriétés, permettant ainsi de les comparer.

2. mise en œuvre pratique, à partir d'implémentations existantes et de nouveaux développements (python, prolog, clingo), pour permettre de les comparer expérimentalement. Cette étape nécessitera d'identifier des formats communs aux différents formalismes.

3. illustration sur des exemples simples, permettant de mettre en évidence le comportement de chaque approche dans des situations concrètes tirées de l'éthique computationnelle et de l'argumentation afin de comparer les résultats dans ces mêmes situations.

Quelques références

Batusov, V., & Soutchanski, M. (2018). Situation Calculus Semantics for Actual Causality. *Proc. of the 22nd AAAI Conf. on Artificial Intelligence, AAAI-18* (pp. 1744–1752).

Beckers, S. (2021). The Counterfactual NESS Definition of Causation. *Proc. of the AAAI Conf. on Artificial Intelligence, AAAI-21* (pp. 6210–6217).

Halpern, J. Y. (2016). *Actual Causality.* The MIT Press.

Kueffner, K. R. (2021). A comprehensive survey of the actual causality literature. Master's thesis, Technische Universität Wien.

Sarmiento, C., Bourgne, G., Inoue, K., & Ganascia, J.-G. (2022). Action Languages Based Actual Causality in Decision Making Contexts. *Proc. of the 24th Int. Conf. on Principles and Practice of Multi-Agent Systems, PRIMA-22* (p. 243–259).