Internship proposal
Root cause analysis based on difference graph discovery

Context: The Pierre Louis Institute of Epidemiology and Public Health (IPLESP), co-accredited by Inserm and Sorbonne University, brings together research strengths in epidemiology and public health within Sorbonne University. IPLESP's main objective is to produce original knowledge on pressing public health issues and related intervention effectiveness, focusing on emerging infectious diseases, chronic diseases, environmental health, and mental health. To tackle these challenges, causal inference [Pearl, 2000, Hernan and Robins, 2023] emerges as an indispensable tool. Therefore, at IPLESP, we are establishing a new team dedicated to developing advanced methodologies rooted in causal inference. One objective of this team is to address root cause analysis challenges in epidemiology, requiring robust methodologies amid increasing data volumes and the intricate interplay of various factors.

Consider a setting in which observational data contain anomalies. Root cause analysis seeks to identify which of these anomalies are the original sources that have causally propagated and triggered the others. This problem has garnered increasing interest in recent years, leading to notable methodological progress [Budhathoki et al., 2022, Ikram et al., Assaad et al., 2023, Zan et al., 2024]. In particular, the method proposed in Assaad et al. [2023] combines prior knowledge encoded in a Summary Causal Graph with time series data from both normal and anomalous regimes to infer a Difference Graph [Assaad, 2025] that reveals changes in causal mechanisms. However, this approach relies on a linearity assumption. More recently, new methods have been developed to learn Difference Graphs directly from data [Wang et al., 2018, Chen et al., 2023, Malik et al., 2024, Bystrova et al., 2024]–without requiring prior causal knowledge or assuming linear relationships. Yet, these data-driven approaches have not been explored in the context of root cause analysis and they have not been extended to time series.

Proposal: The primary objective of this internship is to explore the use of Difference Graph discovery methods for root cause analysis, with a particular focus on time series data. The goal is to design a novel algorithm for discovering Difference Graphs suited to root cause detection, and to assess its effectiveness on real-world epidemiological datasets.

Required skills: Highly motivated candidate with an M2 degree and strong background in probability, machine learning, and causal inference, along with a keen interest in epidemiology. Proficiency in programming is also required. The candidate will have the opportunity to persue a PhD in causal inference.

Location: The intern will work at IPLESP (`https://iplesp.fr/`), located in Paris. She/he will be supervised by Timothée Loranchet, Daria Bystrova, and Charles Assaad.

Dates: Starting date: To be discussed, early 2026, for a duration of 5-6 months.

Contact: To apply, please send a CV and a cover letter to Timothée Loranchet timothee.loranchet@inserm.fr

References

C. K. Assaad. Causal reasoning in difference graphs. In *Proceedings of the Third Conference on Causal Learning and Reasoning*, Proceedings of Machine Learning Research. PMLR, 2025.

C. K. Assaad, I. Ez-Zejjari, and L. Zan. Root cause identification for collective anomalies in time series given an acyclic summary causal graph with loops. In F. Ruiz, J. Dy, and J.-W. van de Meent, editors, *Proceedings of The 26th International Conference on Artificial Intelligence and Statistics*, volume 206 of *Proceedings of Machine Learning Research*, pages 8395–8404. PMLR, 25–27 Apr 2023.

K. Budhathoki, L. Minorics, P. Bloebaum, and D. Janzing. Causal structure-based root cause analysis of outliers. In K. Chaudhuri, S. Jegelka, L. Song, C. Szepesvari, G. Niu, and S. Sabato, editors, *Proceedings of the 39th*

International Conference on Machine Learning, volume 162 of *Proceedings of Machine Learning Research*, pages 2357–2369. PMLR, 17–23 Jul 2022.

D. Bystrova, E. Devijver, V. Manucharian, J. Mondet, and P. Mossuz. Difference graph over two populations: Implicit difference inference algorithm. In *9th Causal Inference Workshop at UAI 2024*, 2024.

T. Chen, K. Bello, B. Aragam, and P. Ravikumar. iscan: identifying causal mechanism shifts among nonlinear additive noise models. In *Proceedings of the 37th International Conference on Neural Information Processing Systems*, NIPS '23, Red Hook, NY, USA, 2023. Curran Associates Inc.

M. Hernan and J. Robins. *Causal Inference: What If*. Chapman & Hall/CRC Monographs on Statistics & Applied Probab. CRC Press, 2023. ISBN 9781420076165.

A. Ikram, S. Chakraborty, S. Mitra, S. Saini, S. Bagchi, and M. Kocaoglu. Root cause analysis of failures in microservices through causal discovery. In S. Koyejo, S. Mohamed, A. Agarwal, D. Belgrave, K. Cho, and A. Oh, editors, *Advances in Neural Information Processing Systems*, volume 35, pages 31158–31170. Curran Associates, Inc.

V. Malik, K. Bello, A. Ghoshal, and J. Honorio. Identifying causal changes between linear structural equation models. In *The 40th Conference on Uncertainty in Artificial Intelligence*, 2024.

J. Pearl. *Causality: Models, Reasoning, and Inference*. Cambridge University Press, New York, NY, USA, 2000. ISBN 0-521-77362-8.

Y. Wang, C. Squires, A. Belyaeva, and C. Uhler. Direct estimation of differences in causal graphs. In *Proceedings of the 32nd International Conference on Neural Information Processing Systems*, NIPS'18, page 3774–3785, Red Hook, NY, USA, 2018. Curran Associates Inc.

L. Zan, C. K. Assaad, E. Devijver, E. Gaussier, and A. Aït-Bachir. On the fly detection of root causes from observed data with application to it systems. In *Proceedings of the 33rd ACM International Conference on Information and Knowledge Management*, CIKM '24, page 5062–5069, New York, NY, USA, 2024. Association for Computing Machinery. ISBN 9798400704369. doi: 10.1145/3627673.3680010.