Post-doctoral study

Causal discovery of extended summary causal graphs for noisy-OR models of event sequences

Context

Networks such as modern telecommunications networks or distributed embedded systems are permanently monitored to allow identification of failure situations; thousands of new data points reflecting the system state changes are generated every minute. Even if faults are rare events, they can easily propagate driven by local and remote dependencies, which makes it challenging to distinguish causes from effects among the thousands of highly correlated alerts.

A timely automated identification and root cause analysis (RCA) of the origins of performance issues allows executing the most adequate corrective actions and preventing their further propagation. In general, RCA is a hard problem, because it requires a deep knowledge of cause-effect dependencies among many features, physical and logical components the network nodes. In a data driven approach, where most of this knowledge is unavailable a priori, a major difficulty emanates from hidden or unknown variables. Furthermore, even in a fully observable system we are faced with the combinatorial explosion of potential cause-effect dependencies and the difficulty to collect enough information for distinguishing causality from spurious correlations.

Goals

The objective of this project is to develop methods to infer causal graphs from observational time series/event-type data generated according to generic noisy-OR models [1]. The causal graphs considered can either be full window causal graphs or a summarized version as extended summary causal graphs [2] and may contain or not hidden common causes. Generic noisy-OR models are structural causal models (SCM) with noisy-OR gates which allow to estimate the effect of multiple causes even if they have never been observed together. We will consider here both simple noisy-OR models in which the noisy-OR gates directly define the SCM, and complex ones in which the noisy-OR gates are sub-parts of an underlying SCM.

Approach

We will first focus on the situation with no hidden common causes and will explore methods to infer extended summary causal graphs both with simple and complex noisy-OR models. In the latter case, we will postulate different underlying SCMs which will have to be both plausible and inferable. Among the discovery methods, we want to consider constraint-based, noise-based and score-based methods, which may be applicable to complex noisy-OR models with underlying probabilistic models, as well as methods based on algorithmic information theory. In a second step, we will study the situation with hidden common causes and explore how to adapt the methods developed in the first step.

To apply Interested candidates should send a complete CV with a list of publications and two reference letters to Armen Aghasaryan armen.aghasaryan@nokia-bell-labs.com, Emilie Devijver emilie.devijver@univ-grenoble-alpes.fr, Eric Gaussier eric.gaussier@imag.fr and Gregor Goessler gregor.goessler@inria.fr. Candidates should be pursuing internationally recognized research in ML/AI, or Information Theory with a strong interest in causal inference and causal reasoning.

Starting date / duration 1 year, starting as soon as possible and no later than October 2023.

Location The work will take place in the Grenoble Computer Science Laboratory, located in Saint-Martin-d'Hères, France.

References

[1] L. Jakovljevic, D. Kostadinov, A. Aghasaryan, T. Palpanas. Towards building a digital twin of complex systems using causal modelling. Complex Networks, 2021.

[2] C. K. Assaad, E. Devijver, E. Gaussier. Discovery of extended summary graphs in time series. Uncertainty in Artificial Intelligence, 2022.