Object-Centric Process Mining for Process Analysis and Operational Support*

Giulia Ruffini[1,*,†]

[1]Department of Computer Science, University of Turin, Turin, Italy

Abstract

Object-centric process mining is a novel paradigm which has been gaining an increasing attention thanks to its potential to overcome limitations of traditional process mining techniques. In this PhD project, we aim at developing novel techniques able to leverage an object-centric paradigm to provide data-driven decision support on running process executions. In particular, we plan to develop novel techniques to a) identify process behaviors which impact the process outcome; b) generate diagnostics for deviant behaviors; and, c) deliver predictions on a KPI of interest at a given moment of a process execution. The manuscript will delve into the research questions and the methodology envisioned for each of these goals.

Keywords

Object-Centric Process Mining, Predictive Process Monitoring, Conformance Checking, Pattern Extraction

1. Introduction

An increasing attention in the field of automated process analysis concerns the Object-Centric Process Mining (OC-PM) perspective, which can better represent the complexity and interconnections existing in business operations compared to classical Process Mining (PM) techniques [1, 2]. First, OC-PM handles multi-dimensional event logs better: by considering different objects rather than a single object as case ID and the other objects as attributes, the convergence and divergence problem can be avoided [2]. OC-PM also considers the interrelationships between different objects, allowing one to grasp information that would otherwise be lost.

Related Works. With the development of object-centric process mining and object-centric event-log (OCEL) formats [2, 3], existing process mining techniques must be adapted as well. Several efforts have already been made in this direction. Contributions made by previous work include: the introduction of object-centric process models (e.g., [2] [4]); techniques to explore process models and features in the OC setting (e.g. [5], or [6]); OC-based approaches to performance analysis [7] and constraint monitoring [8]. However, to the best of our knowledge, only a few studies have addressed important process analysis tasks such as process pattern extraction [9], conformance checking techniques [10], and predictive process monitoring [11, 12] within the OC paradigm. Several open challenges exist for each of these tasks.

ICPM Doctoral Consortium and Demo Track 2023

*Corresponding author.

giulia.ruffini@unito.it (G. Ruffini)

0009-0003-4179-496X (G. Ruffini)

This project aims to contribute to this research area by introducing OC techniques to support the previously mentioned process analysis tasks. More precisely, the project aims at answering the following research questions:

RQ1 How to extract OC process patterns which impact the process outcome?
RQ2 How to asses the conformance of process executions considering OC relations?
RQ3 How to leverage the OC paradigm to enhance the quality of process predictions?

The rest of this manuscript delves into each of these research problems, elaborating upon the corresponding challenges and the envisioned methodology.

2. OC-Pattern Extraction

Research Problem. Process discovery techniques allow one to automatically extract a process model from a given event log tracking past process executions. A well-known limitation of these techniques is that many real-world processes are characterised by complex and variable behaviours, that lead to the extraction of overly complex models which offer little or no support to the analysis. A possible strategy to mitigate this issue consists in inferring interesting process behaviours, rather than start-to-end process model. A process pattern, in classical PM, is a sequence of activities that repeats over multiple cases. Traditional process pattern discovery techniques return process patterns only focused on ordering relations existing between activities occurring in the same process execution. In this way, potential interactions with process objects and with activities occurring in another process instance cannot be captured. Few papers leverage object-centric domain in pattern extraction, e.g. [9], with certain limitations since authors considered flattened event-logs, leading to information loss, and focus on the frequency to measure the importance of a pattern, which often leads to generating a multitude of not-interesting patterns.

Research Goal. We aim at developing novel pattern mining techniques to extract OC process patterns. Following recent approaches for process pattern extraction [13], we aim at extracting patterns that are predictive of process outcome rather than focusing solely on frequent patterns. To this end, we plan to extend previous pattern extraction approaches leveraging graph mining and data abstraction techniques to: a) extend the set of pattern relations considered during the mining to include object-dependent relations; b) develop tailored analysis techniques and tools to allow human users to navigate the set of mined patterns, exploring both intrinsic pattern characteristics and possible relations with other patterns.

Test and Validation. The proposed procedures will be validated using optimisation methods to find particular patterns, e.g. with the highest correlation with the outcome. These techniques will be tested on publicly available datasets after any necessary data pre-processing. Another validation can be performed by considering a dataset of Italian public tenders: after exploring the relational tables and extracting an OCEL from them [14], patterns can be extracted using the investigated techniques.

3. OC-Conformance Checking

Research problem. Identifying anomalies present in an event-log requires a process model and a case notion to compare with it. Classic conformance-checking techniques generate diagnostics which consider process executions in isolation, neglecting possible OC relations and thus potentially missing some deviant behaviours. Adapting conformance checking to the OC paradigm would allow us to overcome this limitation, enabling more accurate diagnostics. Some previous work (e.g., [10] and [5]) proposed techniques to calculate conformance metrics in an OC domain. However, authors worked with "flattened" event-logs.

Research Goal. We aim at developing novel OC conformance checking techniques able to handle the graph-structure of OCEL. In this setting, comparing an individual case with the model entails comparing a graph extracted from OCEL with a graph representing admissible behaviors within the considered model. We plan to investigate and leverage previous work introducing OC case notions [15] and OC process models (e.g., the OC-PN model [2], or the OC-DFG model [4]) to develop techniques able to find a particular morphism to map a graph onto another graph, to identify conforming behaviours and anomalies in OCEL. We also aim at investigating the use of heuristics to mitigate possible computational challenges usually associated with graph analysis techniques.

Test and Validation. To validate the results, we first plan to use a synthetic dataset to test the capability of the proposed method to detect injected anomalies. We then plan to explore publicly available datasets.

4. OC-Predictive Process Monitoring

Research problem. Predictive process monitoring techniques aim at predicting the unfolding of an ongoing process execution to determine, e.g. the next activity to be executed or the total duration of the execution. Previous studies have pointed out the potential benefits of leveraging the OC paradigm for prediction tasks, showing that taking OC relations into account often results in improved prediction performance. As an example, [11] presented a predictive approach using gradient boosting, while [12] proposed another predictive approach using LSTM and neural networks. Despite considering a multi-dimensional event log, both techniques operate with an enriched flattened event log and a multiset of sequences. This approach can lead to a loss of information stemming from the relationships between objects.

Research Goal. We aim at developing novel predictive process monitoring techniques to leverage the graph structure of OC event logs for prediction. To this end, we intend to investigate the use of graph neural networks (GNN), given their ability to deal with graph structures natively. Furthermore, we intend to investigate how to represent additional relations among process activities (e.g., relations between event data, or ordering relations different from the directly following ones) and to develop tailored XAI techniques to analyse the impact of the different relations on the predicted outcome.

Test and Validation. We plan to test and validate the developed techniques on publicly available datasets using traditional classification performance metrics, such as AUC and f1.

References

[1] W. M. P. Van Der Aalst, Process Mining - Data science in action, Second Edition, Springer, 2016.

[2] W. M. P. Van Der Aalst, Object-centric process mining: Dealing with divergence and convergence in event data, in: Software Engineering and Formal Methods: 17th International Conference, SEFM 2019, Oslo, Norway, September 18–20, 2019, Proceedings 17, Springer, 2019, pp. 3–25.

[3] A. F. Ghahfarokhi, G. Park, A. Berti, W. M. P. Van Der Aalst, Ocel: A standard for object-centric event logs, in: European Conference on Advances in Databases and Information Systems, Springer, 2021, pp. 169–175.

[4] W. M. P. Van Der Aalst, A. Berti, Discovering object-centric petri nets, Fundamenta informaticae 175 (2020) 1–40.

[5] A. Berti, W. M. P. Van Der Aalst, Oc-pm: analyzing object-centric event logs and process models, International Journal on Software Tools for Technology Transfer 25 (2023) 1–17.

[6] J. N. Adams, G. Park, S. Levich, D. Schuster, W. M. P. Van Der Aalst, A framework for extracting and encoding features from object-centric event data, in: International Conference on Service-Oriented Computing, Springer, 2022, pp. 36–53.

[7] G. Park, J. N. Adams, W. M. P. Van Der Aalst, Opera: Object-centric performance analysis, in: International Conference on Conceptual Modeling, Springer, 2022, pp. 281–292.

[8] G. Park, W. M. P. Van Der Aalst, Monitoring constraints in business processes using object-centric constraint graphs, in: International Conference on Process Mining, Springer, 2022, pp. 479–492.

[9] E. R. Faria Junior, T. R. Neubauer, M. Fantinato, S. M. Peres, Clustering analysis and frequent pattern mining for process profile analysis: An exploratory study for object-centric event logs, in: International Conference on Process Mining, Springer, 2022, pp. 269–281.

[10] J. N. Adams, W. M. P. Van Der Aalst, Precision and fitness in object-centric process mining, in: 2021 3rd International Conference on Process Mining (ICPM), IEEE, 2021, pp. 128–135.

[11] R. Galanti, M. De Leoni, N. Navarin, A. Marazzi, Object-centric process predictive analytics, Expert Systems with Applications 213 (2023) 119173.

[12] W. Gherissi, J. El Haddad, D. Grigori, Object-centric predictive process monitoring, in: International Conference on Service-Oriented Computing, Springer, 2022, pp. 27–39.

[13] V. Mozhgan, L. Genga, X. Lu, R. Verhoeven, H. van Laarhoven, R. Dijkman, Interactive multi-interest process pattern discovery (2023).

[14] A. Berti, G. Park, M. Rafiei, W. M. P. Van Der Aalst, A generic approach to extract object-centric event data from relational databases (2023).

[15] J. N. Adams, D. Schuster, S. Schmitz, G. Schuh, W. M. P. Van Der Aalst, Defining cases and variants for object-centric event data, in: 2022 4th International Conference on Process Mining (ICPM), IEEE, 2022, pp. 128–135.