Multi-Dimensional Decision Mining: Discovering Decision Rules in Object-Centric Event Logs

Erfan Elhami[1,2]

[1]School of EECS, University of Ottawa, Ottawa, Ontario, Canada.
[2]Telfer School of Management, University of Ottawa, Ottawa, Ontario, Canada.

Abstract

Decision Mining (DM) aims to explain process behavior by identifying the decision logic underlying routing choices. However, conventional DM approaches rely on case-centric event logs, meaning they cannot fully capture the complex multi-object interactions inherent in real-world business processes. In this thesis, I introduce a novel DM framework for object-centric processes, which exploits the object relations present in object-centric event logs (OCEL), while maintaining the interpretability and expressivity of the discovered decision rules. This study follows the Design Science Research methodology. Early work included a systematic literature review, accompanied by an ontology that formally captures the various aspects of DM. This initial work enabled the identification of gaps and research opportunities in object-centric DM. Currently, multiple potential solution approaches are being investigated, including inductive logic programming, graph-based learning, and knowledge distillation. Subsequent phases will implement the framework and evaluate it empirically on synthetic and real-world OCEL datasets.

Keywords

Decision Mining, Decision Point Analysis, Process Discovery, Object-Centric Process Mining

1. Introduction

Process Discovery is a well-developed area of *Process Mining (PM)* that aims to identify a process's behavior from event logs [1]. Based on sequences of events, it discovers a high-level representation of process behavior, which enables identification of workflow variants, behavioral patterns, and alternative process paths. Yet, it fails to explain the underlying reasons for the observed behavior—i.e., why objects follow particular paths through the process [2, 3]. *Decision Mining (DM)* is a subfield of PM that identifies criteria for choosing branches at *Decision Points (DP)*. For example, at a triage DP, DM could determine that "*oxygen saturation* below 90%" is a criterium for following the «Admission» branch. Rozinat et al. [2] proposed the concept of DM in 2006 by describing each DP as a classification problem. Typically, DM incorporates Machine Learning (ML) algorithms to derive decision rules.

In parallel, *Object-Centric Process Mining (OCPM)* has emerged as a novel PM paradigm that captures the multi-object interactions inherent in processes [4, 5]. Compared to a conventional *case-centric event log*, where each event is tied to a single case identifier, an *object-centric event log (OCEL)* allows an event to be linked to multiple objects simultaneously [4, 6]. For example, a surgery process simultaneously involves a *patient*, several *physicians*, and *medical supplies*. To explain behavior at a particular point in the process, information (such as events) related to all these objects can thus be exploited [4]. Decisions rarely depend on a single isolated case, but rather on shared resource constraints, physician workloads, and evolutions of object attributes [4]. Thus, it is vital to capture information across all these interacting entities. In this thesis, I define the concurrent temporal evolution of object lifecycles alongside their structural interdependencies—i.e., direct Object-to-Object (O2O) and Event-to-Object (E2O) relationships—as multi-dimensional dynamics. To capture multi-dimensional dynamics, OCELs inherently utilize a heterogeneous graph design—comprising distinct node types for events and objects,

Proceedings of the BPM 2026 Best Dissertation Award, Doctoral Consortium, and Demonstration & Resources Forum (BPM 2026), Toronto, Canada, September 27th to October 2nd, 2026.
erfan.elhami@uottawa.ca (E. Elhami)
0009-0003-0232-3209 (E. Elhami)

and edges representing O2O and E2O relations—thereby preserving object-relational information that remains hidden in case-centric logs [6, 7].

It has been argued that decision-making in real-world processes is also object-centric [8]. For instance, deciding a patient's treatment pathway may depend not only on the patient object's clinical priority, but also on hospital resource objects, such as physician expertise and equipment availability. Most research to date has been limited to case-centric DM, which recognizes *process-spanning* and *instance-spanning* decisions [9, 10]; for instance, where a decision on an «ICU Admission» depends not only on a patient's condition, but also on instance-spanning priorities across concurrent cases and process-spanning resource constraints (e.g., physician workload and equipment availability). Earlier non-OCEL techniques attempted to incorporate such context by injecting related data into the event log, for instance, as discrete attributes [3]. In general, however, case-centric event logs fundamentally lack the structure required to capture this information natively [3, 4]. Such approaches thus ignore the multi-dimensional dynamics found in OCELs as described earlier. Object-centric DM introduces a multi-dimensional perspective for identifying complex decision rules driven by cross-object interdependencies [3, 8]. Despite this potential, object-centric DM remains a less-explored area in PM research [8]. This gap presents a significant research opportunity: to design and develop a novel DM framework that directly leverages the complex relational information in OCEL to identify the distributed decision logic among interacting process objects [3, 8].

2. State of the Art

Conventional DM considers each DP as a classification problem. First, related data for each DP is collected and transformed into an *attribute–value structure*, where each *row* represents a case, and each *column* represents a feature describing the state upon reaching the DP. Complex types of information (e.g., time series data) are usually converted into discrete features and integrated into the tabular representation [11, 12]. Each case is then labeled according to the associated choice, that is, the chosen DP branch. The labeled data is subsequently employed to train an interpretable classification model (typically a decision tree) from which decision rules can be derived. Collectively, these rules explicitly represent the conditions that influence the selection of alternative process paths. Thus, conventional DM requires transforming all relevant process and contextual data into a case-based tabular dataset. Generally, reducing graph-structured data to a flattened, bi-dimensional representation has several significant limitations, including the imposition of an artificial linear ordering, an inability to capture dependencies and hidden relational semantics, and a heavy reliance on inflexible, hand-engineered features [13].

To the best of my knowledge, the *Integrated Object-centric Decision Discovery Algorithm (IODDA)* introduced by Goossens et al. [8] is the only current work on object-centric DM. This approach builds on and extends the case-centric *Variable-Activity Pair (VAP)* method by De Smedt et al. [14] . Notably, this work extracts decision rules by tracking how activity executions change process variables (i.e., *variable shifts*), rather than by analyzing a predetermined DP. The authors aimed to adopt this algorithm for a multi-object context without flattening the event log, and thus avoiding the effects of *convergence* (i.e., artificial duplication of events [7]) and *divergence* (i.e., loss of explicit object-instance assignments [7]). To that end, they utilized a native multi-dimensional data structure termed *Data-Aware Object-Centric Event Log (DOCEL)* [15]. Instead of the aforementioned bi-dimensional pairs, IODDA extracts *Attribute-Activity-Object Type Shifts (AAOTS)*, tracking attribute changes relative to multi-object lifecycles rather than a case identifier. To capture correlations among different objects involved in a decision, the algorithm cross-analyzes intersecting event paths across object traces. For instance, in the earlier «ICU Admission» example, the decision logic could be dependent upon the attributes of both the *Patient* (e.g., severity score) and the *Physician* (e.g., previous decisions). Consequently, IODDA produces a *holistic Decision Requirements Diagram (DRD)* that maps variables to their respective object types.

This approach has two critical limitations. First, at the data preparation layer, the DOCEL structure was developed to formalize variable states and their changes with respect to object lifecycles, thereby avoiding the OCEL 1.0 limitations in tracking dynamic object attributes [15]. The recent introduction of the OCEL 2.0 standard natively incorporates *Evolving Object Attribute Values* [16], thereby removing this requirement for intermediate formats. Moreover, because the DOCEL metamodel relies on explicit *E2O* foreign keys, it can neither capture standalone *O2O* relationships independent of events, nor process qualified semantic roles (i.e., *Qualifiers*) natively supported in OCEL 2.0. Instead, IODDA heuristically extracts O2O dependencies from overlapping object traces rather than exploiting direct semantic relations. As in the aforementioned «ICU Admission» example, the decision logic may depend on a physician's role as a *Primary Surgeon* or an *Assistant* (i.e., a qualifier), alongside a direct O2O link between a *specialized monitoring device* and the *ICU department.* As DOCEL lacks native qualifiers and relies solely on E2O links, IODDA treats all involved physicians identically and can only artificially infer O2O relations from overlapping events. Consequently, the extracted rules may become overgeneralized and fail to capture the exact decision criteria. Second, this pipeline relies on human domain experts, who must manually configure parameter thresholds (e.g., minShift and minCorr) and filter out spurious statistical correlations that lack substantive business semantics. These limitations point towards a strong research opportunity to develop a native DM method that operates directly upon a standard OCEL 2.0 relational structure.

3. Problem Statement and Research Questions

The following deficiencies have been recognized:

Challenge 1: Relational Blindness and Flattening Limitations: In OCEL 2.0, relationships are no longer restricted to event-based connections (i.e., E2O relationships); logs natively include direct *O2O relationships* (e.g., a connection between a doctor and a hospital), *dynamic attributes* (e.g., a doctor's evolving certifications), and relationship *qualifiers* [16]. Conventional DM methods do not consider these rich, event-independent relationships. To interpret them with commonly used ML classifiers (e.g., decision trees), researchers must flatten the log into attribute-value tables. This causes a significant limitation as it reintroduces the dimension-reduction problems that originally motivated the switch away from case-centric logs [8].

Challenge 2: Temporal Evolution and Multi-Object History Modeling: Given that real-world processes are highly dynamic and that object attributes change constantly over time, decisions often depend on a combination of concurrent object conditions and their previous transaction histories [3, 9, 10, 11, 12, 17]. For example, when a doctor evaluates a patient at a DP to choose between «ICU Admission» and «Surgery», this path choice does not depend solely on a static snapshot of the patient, but also on the temporal evolution of their condition (e.g., changing vital signs over the past six hours) and the cumulative historical profiles of all interacting objects up to that point (e.g., the availability of the surgery room, and number of hours that the doctor has already spent in surgery). A substantial challenge is that conventional DM models are not designed to capture such temporal evolution, nor do they consider the independent historical profile of objects before they intersect at a DP [3].

Challenge 3: Dependency on Domain Experts and Manual Preprocessing: Previous work on object-centric DM requires human domain experts; they rely on statistical correlation parameters that require specific hyperparameter adjustments, such as defining precise thresholds for minimum variable shifts [8]. Moreover, real-life event logs often include attributes with non-scalar, structured formats, such as nested lists or dictionary elements (e.g., a patient's treatment cost attribute containing a list of values of physician fees, medication charges, and facility surcharges) that specify feature variations. Although current DM approaches through feature engineering could extract object relations or flatten such complex schemas for standard classifiers, they require extensive manual preprocessing and feature normalization [8]. Such dependence on human expertise challenges the automated functionality of DM.

In light of this, the following research questions are presented:

- **RQ1 (Object History and Temporality):** How can the continuous temporal evolution and historical lifecycles of interacting process objects be leveraged to discover decision rules that capture their cumulative behavior over time?

- **RQ2 (OCEL 2.0 Relational Richness and Complex Attributes):** How can an object-centric DM framework natively leverage the relational topology and rich information of the OCEL 2.0 standard to discover realistic, context-aware business decision rules? Note: OCEL 2.0 information includes comprehensive multi-dimensional relationships, relationship qualifiers, evolving attribute values, and complex non-scalar formats.

4. Research Methodology

This research adopts the *Design Science Research (DSR)* paradigm [18, 19] to address the research gaps highlighted in Section 3. The research objective is to develop and evaluate a novel DM framework for object-centric event data (main artifact). The study follows the DSR procedure through the stages of *Problem Identification*, *Objective Definition*, *Artifact Design and Development*, and *Evaluation*. The operational feasibility of the main artifact will be evaluated using both real-world and synthetic OCEL datasets, in addition to domain expert evaluations, in accordance with the DSR paradigm.

5. Proposed Research Directions

This thesis proposes a novel DM framework leveraging the native relational topology of OCEL 2.0 and objects' lifecycles to produce interpretable decision rules. The framework is designed around four main stages: (i) identifying decision-relevant contexts by extracting multi-dimensional relationships, relationship qualifiers, evolving attribute values, and objects' temporal histories from OCEL (addressing Challenge 1, Challenge 2, and RQ1); (ii) representing this heterogeneous graph and temporal information without inducing dimension-reduction or flattening distortions (addressing Challenge 1 and RQ2); (iii) discovering interpretable, context-aware business decision rules from the enriched representation through automated learning (addressing Challenge 3 and RQ2); and (iv) evaluating the quality of the discovered rules according to predefined criteria. To address stages (ii) and (iii), arguably the most challenging stages, I plan to pursue the following directions:

Inductive Logic Programming (ILP): ILP is an ML technique that learns first-order logic rules directly from multi-relational graphs and domain background knowledge (BK) [20, 21]. Unlike common classification techniques that flatten graph data, ILP directly manipulates first-order clauses, hence preserving the relational and topological integrity of the log [20, 22, 23]. Modern, state-of-the-art ILP systems (e.g., Popper) can induce optimal, recursive logic programs by searching the hypothesis space via a generate-test-and-constrain loop [21]. Natively incorporating BK allows the framework to encode historical event sequences, evolving attribute values, and object lifecycle transitions. This support for recursion and first-order relational reasoning contributes to RQ1, providing the formal modeling foundation needed to capture the temporal history of interacting objects and to extract human-readable, context-aware decision rules without the dimensional reduction or distortions of flattened datasets.

Graph-based Learning: By encoding OCEL data as graphs, graph-based ML techniques such as *Graph Neural Networks (GNNs)* can be applied. Graph-based learning can directly learn from relationships among events and objects, thereby exploiting the topological characteristics of object-centric processes [13, 24]. Such models could represent complex dependencies, including relationships among object types and object lifecycle interdependencies, which are difficult to model using conventional tabular representations [4, 24]. For graph-based learning, OCEL can be transformed into suitable structures, such as *Event Knowledge Graphs* [25], or represented directly in native graph formats (e.g., HOEG [24]). However, a significant obstacle here is that GNN-based models are often black-box, which may limit interpretability and the extraction of explicit decision rules [4, 13]. This issue motivates the subsequent direction.

Knowledge Distillation (KD): The third direction is to apply KD, where knowledge acquired by an advanced *teacher model* (e.g., a GNN analyzing OCEL) is transferred to a simpler, interpretable *student model* (e.g., decision tree) [26, 27]. This approach may facilitate the use of rich relational information in OCEL through graph-based learning while maintaining the required interpretability of decision rules [27, 28]. Developed frameworks, such as *GraphChef* [29] and *Tree-Network-Tree* [26], and logical distillation via *Iterated Decision Trees* (IDT) [30], suggest the viability of extracting knowledge from graph models to create interpretable learners, such as decision trees. Moreover, the IDT approach extracts symbolic rules in first-order logic, thereby creating a connection between GNNs and ILP techniques [30]. Although proving that the student model accurately retains decision-relevant knowledge remains a significant challenge [26, 27, 29], combining graph learning with symbolic rule extraction primarily addresses RQ2, enabling the framework to exploit OCEL 2.0's full relational richness while extracting interpretable, context-aware decision rules.

6. Research Roadmap and Current Progress

The research completed the DSR phases (introduced in Section 4) of *problem identification* and *objective definition.* A *systematic literature review (SLR)* [3] was conducted to identify the previous advancements in DM and to highlight existing research gaps. The results of the review and DM concepts were formalized by developing an *ontology* for the DM research [31]. Based on gap analysis, DM in object-centric processes has been studied, including a detailed analysis of Goossens et al. [8] , to identify current advancements and open research challenges. Several candidate solution directions have been identified. The next DSR phases will be centered on *artifact design*, *development*, and *evaluation.*

An *empirical evaluation* methodology will be used to assess the proposed framework. The overall quantitative capabilities of the framework will be evaluated using k-fold cross-validation on publicly available event logs in the standard OCEL 2.0 format (such as Logistics or Order Management logs from the official repository [32]). The performance of the proposed framework will be evaluated against the *IODDA algorithm* by Goossens et al. [8] , the only object-centric DM framework (at the time of writing); and traditional *case-centric DM* approaches, including those focusing on process-spanning constraints. I will use standard statistical metrics (*Accuracy*, *Precision*, *Recall*, and *F1-score*). The discovered decision rules will also be evaluated individually against rule-based criteria – specifically *Interpretability*, *Effective Complexity*, *Parsimony*, and *Remine Consistenc*y – in line with Wais et al. [33] . In addition, for a real-world dataset, the discovered rules will be validated by a domain expert, whereas for a synthetic log, they will be evaluated against pre-designed ground-truth rules injected during log synthesis.

7. Conclusions

Existing DM approaches have mainly been developed for case-centric event logs, while limited research has investigated how the relational information available in OCEL can be leveraged for DM. This research seeks to exploit OCEL's object-relational information while preserving the interpretability and expressivity of the discovered decision rules. This study follows the DSR methodology. In the initial stages, an SLR and an ontology have been developed to explore the state of the art and identify current limitations of DM in object-centric processes. Future work will concentrate on the design, development, and evaluation of a novel DM framework that leverages OCEL's object-relational information to generate realistic, interpretable decision rules.

Declaration on Generative AI

The author used AI-assisted tools (e.g., Grammarly and Claude) for grammar and spelling checks, paraphrasing, text refinement, and critique. No content or text was created through generative AI.

References

[1] W. M. P. van der Aalst, J. Carmona, Process Mining Handbook, Springer Nature, 2022. doi:10.1007/978-3-031-08848-3.

[2] A. Rozinat, W. M. P. van der Aalst, Decision mining in business processes, Technische Universiteit Eindhoven, Eindhoven, 2006. URL: https://research.tue.nl/files/1750830/609514.pdf.
[3] E. Elhami, W. Van Woensel, D. Amyot, Discovering the why of process models – a systematic literature review on decision mining, SSRN, 2026. doi:10.2139/ssrn.6749580.
[4] A. Seidel, M. Weske, M. Montali, et al., Object-centric process management: A research manifesto, Information Systems 141 (2026) 102728. doi:10.1016/j.is.2026.102728.
[5] W. M. P. van der Aalst, Object-centric process mining: Dealing with divergence and convergence in event data, in: P. C. Ölveczky, G. Salaün (Eds.), Software Engineering and Formal Methods, Springer, 2019, pp. 3–25. doi:10.1007/978-3-030-30446-1_1.
[6] J. N. Adams, D. Schuster, S. Schmitz, G. Schuh, W. M. van der Aalst, Defining cases and variants for object-centric event data, in: 2022 4th International Conference on Process Mining (ICPM), 2022, pp. 128–135. doi:10.1109/ICPM57379.2022.9980730.
[7] W. M. P. van der Aalst, Object-centric process mining: Unraveling the fabric of real processes, Mathematics 11 (2023). doi:10.3390/math11122691.
[8] A. Goossens, J. De Smedt, J. Vanthienen, Extracting process-aware decision models from object-centric process data, Information Sciences 682 (2024) 121263. doi:10.1016/j.ins.2024.121263.
[9] K. Winter, F. Stertz, S. Rinderle-Ma, Discovering instance and process spanning constraints from process execution logs, Information Systems 89 (2020) 101484. doi:10.1016/j.is.2019.101484.
[10] K. Winter, S. Rinderle-Ma, Discovering instance-spanning constraints from process execution logs based on classification techniques, in: 2017 IEEE 21st International Enterprise Distributed Object Computing Conference (EDOC), IEEE CS, 2017, pp. 79–88. doi:10.1109/EDOC.2017.20.
[11] B. Scheibel, S. Rinderle-Ma, Decision mining with time series data based on automatic feature generation, in: Advanced Information Systems Engineering, Springer, 2022, pp. 3–18. doi:10.1007/978-3-031-07472-1_1.
[12] R. Dunkl, S. Rinderle-Ma, W. Grossmann, K. Anton Fröschl, A method for analyzing time series data in process mining: Application and extension of decision point analysis, in: Information Systems Engineering in Complex Environments, Springer, 2015, pp. 68–84. doi:10.1007/978-3-319-19270-3_5.
[13] J. Zhou, G. Cui, S. Hu, Z. Zhang, C. Yang, Z. Liu, L. Wang, C. Li, M. Sun, Graph neural networks: A review of methods and applications, AI Open 1 (2020) 57–81. doi:10.1016/j.aiopen.2021.01.001.
[14] J. De Smedt, F. Hasić, S. K. L. M. vanden Broucke, J. Vanthienen, Holistic discovery of decision models from process execution data, Knowl Based Syst 183 (2019) 104866. doi:10.1016/j.knosys.2019.104866.
[15] A. Goossens, J. De Smedt, J. Vanthienen, W. M. P. van der Aalst, Enhancing data-awareness of object-centric event logs, in: Process Mining Workshops, Springer, 2023, pp. 18–30. doi:10.1007/978-3-031-27815-0_2.
[16] A. Berti, I. Koren, J. N. Adams, G. Park, B. Knopp, N. Graves, M. Rafiei, L. Liß, L. T. G. Unterberg, Y. Zhang, C. Schwanen, M. Pegoraro, W. M. P. van der Aalst, Ocel (object-centric event log) 2.0 specification, 2024. URL: https://arxiv.org/abs/2403.01975.
[17] A. Banham, S. J. J. Leemans, M. T. Wynn, R. Andrews, xPM: A framework for process mining with exogenous data, in: J. Munoz-Gama, X. Lu (Eds.), Process Mining Workshops, Springer, 2022, pp. 85–97. doi:10.1007/978-3-030-98581-3_7.
[18] A. R. Hevner, S. T. March, J. Park, S. Ram, Design science in information systems research, MIS Quarterly 28 (2004) 75–105. URL: http://www.jstor.org/stable/25148625.
[19] K. Peffers, T. Tuunanen, M. A. Rothenberger, S. Chatterjee, A design science research methodology for information systems research, Journal of Management Information Systems 24 (2007) 45–77. doi:10.2753/MIS0742-1222240302.
[20] M. V. M. França, G. Zaverucha, A. S. d'Avila Garcez, Fast relational learning using bottom clause propositionalization with artificial neural networks, Machine Learning 94 (2014) 81–104. doi:10.1007/s10994-013-5392-1.
[21] A. Cropper, R. Morel, Learning programs by learning from failures, Machine Learning 110 (2021)

801–856. doi:10.1007/s10994-020-05934-z.
[22] N. Lavrač, B. Škrlj, M. Robnik-Šikonja, Propositionalization and embeddings: two sides of the same coin, Machine Learning 109 (2020) 1465–1507. doi:10.1007/s10994-020-05890-8.
[23] F. Železný, N. Lavrač, Propositionalization-based relational subgroup discovery with RSD, Machine Learning 62 (2006) 33–63. doi:10.1007/s10994-006-5834-0.
[24] T. K. Smit, H. A. Reijers, X. Lu, HOEG: A new approach for object-centric predictive process monitoring, in: Advanced Information Systems Engineering, Springer, 2024, pp. 231–247. doi:10.1007/978-3-031-61057-8_14.
[25] A. Swevels, D. Fahland, M. Montali, Implementing object-centric event data models in event knowledge graphs, in: Process Mining Workshops, Springer, 2024, pp. 431–443. doi:10.1007/978-3-031-56107-8_33.
[26] J. Li, Y. Li, X. Xiang, S.-T. Xia, S. Dong, Y. Cai, TNT: An interpretable tree-network-tree learning framework using knowledge distillation, Entropy 22 (2020). doi:10.3390/e22111203.
[27] Y. Tian, S. Pei, X. Zhang, C. Zhang, N. V. Chawla, Knowledge distillation on graphs: A survey, ACM Comput. Surv. 57 (2025). doi:10.1145/3711121.
[28] Y. Li, L. Liu, G. Wang, Y. Du, P. Chen, EGNN: Constructing explainable graph neural networks via knowledge distillation, Knowledge-Based Systems 241 (2022) 108345. doi:10.1016/j.knosys.2022.108345.
[29] P. Müller, L. Faber, K. Martinkus, R. Wattenhofer, GraphChef: Decision-tree recipes to explain graph neural networks, in: The Twelfth International Conference on Learning Representations, 2024, pp. 5275–5303. URL: https://openreview.net/forum?id=IjMUGuUmBI.
[30] A. Pluska, P. Welke, T. Gärtner, S. Malhotra, Logical distillation of graph neural networks, 2024. URL: https://arxiv.org/abs/2406.07126. arXiv:2406.07126.
[31] E. Elhami, W. Van Woensel, D. Amyot, Exploring decision mining research: A systematic review report and knowledge graph, in: Business Process Management Workshops, Springer, 2026, pp. 82–93. doi:10.1007/978-3-032-13426-4_7.
[32] OCEL Standard, Event Logs - OCEL 2.0, https://www.ocel-standard.org/event-logs/overview/, 2024. Accessed: 2026-06-07.
[33] B. Wais, S. Rinderle-Ma, Towards a comprehensive evaluation of decision rules and decision mining algorithms beyond accuracy, in: Advanced Information Systems Engineering, Springer, 2024, pp. 403–419. doi:10.1007/978-3-031-61057-8_24.