Business Process Intervention Discovery and Decision Support

Jakob De Moor[1]

[1]Information Systems Engineering Research Group (LIRIS), KU Leuven, Naamsestraat 69, Leuven, Belgium

Abstract

Prescriptive Process Monitoring (PresPM) leverages machine learning to optimize business process outcomes by recommending real-time interventions. Two common approaches for recommending interventions in PresPM are Causal Inference (CI) and Reinforcement Learning (RL). Despite progress in these fields, significant gaps remain, including the lack of comprehensive comparisons of PresPM methods, the assumption of predefined intervention points and dimensions, limited handling of complex interventions (e.g., sequential interventions), and insufficient exploration of process interdependencies. This research proposes addressing these challenges by designing a simulator that generates synthetic data mimicking complex business processes, enabling the evaluation of approaches across diverse intervention types. Additionally, a novel causal decision point discovery algorithm will be developed, tailored for PresPM in business settings. The study also aims to enhance CI and RL techniques for safe and efficient handling of sequential interventions, decision timing, and process interdependencies. These contributions will be validated using both synthetic and real-world datasets, advancing the field of PresPM by providing novel methodologies for complex optimization problems.

Keywords

Prescriptive Process Monitoring, Process Optimization, Causal Discovery

1. Introduction and background

AI-powered systems in Prescriptive Process Monitoring (PresPM) offer transformative opportunities for businesses by enhancing decision-making. PresPM uses machine learning to analyze event log data from business processes, uncovering relationships between process variables and targets like delivery times, loan acceptance, patient recovery, product availability, or default rates. These systems try to optimize outcomes by prescribing real-time interventions, such as machine maintenance, customer calls, or loan application cancellations. Benefits include improved product and service quality, cost savings, and increased staff satisfaction.

After a thorough review of the literature, it is evident that two approaches for recommending interventions are generally adopted: Causal Inference (CI) and Reinforcement Learning (RL). CI comprises two primary subfields. The first, known as Causal Discovery (CD), focuses on identifying causal relationships between variables. Its goal is to uncover the true causal structure of the data, typically represented as a graph. An algorithm is said to solve the CD problem if it converges to the true structure as the sample size increases. In Process Mining (PM), CD techniques are often used to pinpoint decision points within a business process. The second subfield is concerned with estimating the effect of a treatment, often denoted by the Individual Treatment Effect (ITE) in a particular data instance. Current CI approaches in PresPM primarily concentrate on this second subfield, which can be divided into two categories: direct and indirect CI. Direct CI entails training a machine learning model to directly estimate the ITE of an intervention [1, 2, 3]. Indirect CI involves training models to predict the next executed activities for each possible action in the intervention. The ITE is then inferred based on the obtained outcomes of the activity predictions [4]. In general, CI approaches are originally designed for non-sequential intervention optimization problems. Reinforcement Learning (RL) is the other widely adopted approach in PresPM, where an RL agent learns an optimal policy by interacting with an environment and receiving rewards. While RL has shown promising results in controlled settings [2, 5],

ICPM 2024 Doctoral Consortium, October 14–18, 2024, Kongens Lyngby, Denmark
jakob.demoor@kuleuven.be (J. D. Moor)
0009-0006-4788-5346 (J. D. Moor)

real-world online training and implementation poses risks due to its exploratory nature. For example, a bank would avoid using an RL agent for real-time loan approvals because the agent might make risky or random decisions during its exploration phase.

Despite the establishment of these initial approaches, many limitations in PresPM remain. Firstly, there is a lack of extensive comparisons of PresPM methodologies. To the best of my knowledge, no study has thoroughly and comprehensively compared multiple approaches. Additionally, existing methods are typically focused on a narrow range of intervention types and limited complexity, often neglecting more complex optimization problems. As described in [2], interventions can be classified by action width (scope of intervention actions) and action depth (timing and sequence of the intervention). Most research handles interventions with binary action width and fixed action depth [1, 3, 6], though some consider multi-class action width [4] or non-fixed action depth [2]. Sequential interventions remain unexplored, and no approach has been tested or compared across multiple intervention types. A second major issue concerns the inaccuracy in method evaluation. Current research predominantly relies on real-life offline data [1, 3, 4, 6, 7], which can be incomplete due to data-gathering policies, such as a bank's loan application policy, thereby omitting parts of the optimization state space. Evaluating a model's policy on a case-by-case basis requires knowledge of the outcomes under the prescribed actions, which is infeasible with only historical data as it lacks the counterfactual outcomes, i.e., the outcome under an alternative action than the one recorded. Thirdly, most PresPM methodologies assume predefined intervention points. The CD aspect of CI, essential for identifying these points, is overlooked in PresPM, despite extensive research in the PM and CI fields [8, 9, 10, 11]. Understanding data relations in a PresPM system is crucial for identifying interventions and their dimensions, and selecting appropriate prescriptive techniques. The last research gap concerns process interdependencies. Real-world process executions frequently exhibit interdependencies, where one execution influences another. For instance, in a manufacturing facility, if one production process requires a large portion of the available machinery, it can limit the efficiency of other processes occurring simultaneously. Despite the prevalence of these relationships, a limited number of PresPM papers address them [3, 5].

2. Research Objectives and Methodology

To address the identified gaps in the research section, I propose the following research objectives:

2.1. Design and implement a simulator for PresPM

Current PresPM research lacks comprehensive comparisons because it omits complex interventions and a wide range of evaluation perspectives. To address this, I propose a synthetic data generator in Python. Building on [2], the generator will produce data mimicking a loan application process in a bank. It will be able to calculate target variables for every possible action of an intervention, addressing the limitations of historical offline data that lack counterfactual outcomes. The generator will support comparisons across different intervention types and complexities, expanding PresPM research to more complex optimization problems (i.e., intervention sequences). It will also facilitate experimentation within each intervention type by varying generating parameters. To validate the generator, I will implement and compare both CI and RL approaches, along with benchmarking existing methods. This will allow me to determine whether the simulator can effectively and accurately highlight the differences between methods for varying levels of complexity. It will mark the first time a sequential intervention is addressed and the first comprehensive comparison of existing PresPM implementations. A key challenge will be balancing the need to maintain variability and complexity similar to real-life event logs, while also keeping the simulated loan application process simple enough for easy interpretation and method comparison. This balance may involve trade-offs. Although applying methods to simulated data will likely only approximate their performance on real-life logs, the simulator's ability to accurately evaluate method performance for different intervention decision problems makes it a valuable tool for PresPM research.

Although simulators in process mining have seen significant advancements in research [12, 13, 14], our simulator will be uniquely designed for testing and developing PresPM approaches. It will focus on key aspects such as (multiple) intervention types, precise evaluation, and the flexibility to customize parameters for experimentation. Additionally, it will support the implementation of both online methods (e.g., RL) and offline methods (e.g., CI), enabling the generation of both real-time environments and offline datasets.

2.2. Develop a causal decision point discovery algorithm for Prescriptive Process Monitoring in business settings

Current prescriptive approaches often assume predetermined intervention points and dimensions, lacking a systematic methodology for identification. To address this, I will explore and tailor CD methods for PresPM in business settings. One research avenue is the exploration of various PM techniques, such as methods that incorporate structural equation models as described in [8], upper-bound causal graphs as discussed in [9], uplift trees as researched in [15], and the P-MInD framework [11]. To adapt these techniques for use with PresPM, it is essential to enhance them by incorporating the identification of intervention point dimensions (action width and action depth). These dimensions reflect the complexity of the problem, and understanding them is essential for selecting the most appropriate optimization method. I will additionally investigate methods to incorporate expert knowledge to improve CD. The evaluation will include two components: synthetic data, where the true causal structure is known, and real-world datasets to assess practical applicability.

2.3. Develop novel PresPM algorithms, enhancing CI and RL

Current PresPM methodologies overlook complex intervention problems, such as sequential optimization and process interdependencies. I aim to adapt PresPM approaches to address these advanced Machine Learning tasks.

First, I aim to extend CI to handle intricate sequential tasks, considering challenges like decision timing. This includes exploring optimal threshold optimization algorithms [2] and sequence prediction techniques for prescribing a series of decisions. Current research in CI beyond PresPM includes methods for sequential decision-making, such as those proposed by [16]. However, these approaches are still in their early stages and need to be adapted to handle more complex business processes, which often feature multiple execution paths, loops, and parallel activities. To address process interdependencies, I will explore the integration of CI with enhanced heuristics [3] and Object-Centric PM to analyze the behavior of individual objects [17]. I will investigate whether an object-centric perspective can improve the performance of PresPM methods.

Additionally, I will adapt RL to these complex optimization challenges, e.g., intervention sequences and process interdependencies, by improving its training efficiency. I will limit the RL action space to decision points identified by my CD techniques and explore efficient sampling methods [18]. The aim is to evaluate the suitability and effectiveness of these techniques for optimizing complex business processes. Incorporating heuristics in RL will also be investigated to manage the increased search space. To ensure real-world applicability of these adaptations, I will mitigate risks in online RL training by exploring offline RL techniques. These include using process discovery to construct training environments [7], constraining RL to known state spaces using nearest neighbor approaches, and leveraging predictive models for anticipating next states [19]. The objective is to determine which techniques for building offline environments are best suited to different types of business process data and optimization challenges. I will also identify low-risk application domains and implement predefined rules to prevent excessively risky actions by the RL agent.

The novel methods will first be evaluated using the previously developed simulator that can generate all potential outcomes of a case, allowing for a thorough and accurate assessment. Afterward, the practical applicability of these methods will be validated using real-world datasets, ideally across various diverse business settings.

References

[1] Z. Dasht Bozorgi, I. Teinemaa, M. Dumas, M. La Rosa, A. Polyvyanyy, Prescriptive process monitoring based on causal effect estimation, Information Systems (2023). doi:https://doi.org/10.1016/j.is.2023.102198.

[2] H. Weytjens, W. Verbeke, J. De Weerdt, Timed process interventions: Causal inference vs. reinforcement learning, in: J. De Weerdt, L. Pufahl (Eds.), Business Process Management Workshops, Springer Nature Switzerland, Cham, 2024, pp. 245–258.

[3] M. Shoush, M. Dumas, When to intervene? prescriptive process monitoring under uncertainty and resource constraints, in: C. Di Ciccio, R. Dijkman, A. del Río Ortega, S. Rinderle-Ma (Eds.), Business Process Management Forum, Springer International Publishing, Cham, 2022, pp. 207–223.

[4] S. Weinzierl, S. Dunzer, S. Zilker, M. Matzner, Prescriptive business process monitoring for recommending next best actions, in: D. Fahland, C. Ghidini, J. Becker, M. Dumas (Eds.), Business Process Management Forum, Springer International Publishing, Cham, 2020, pp. 193–209.

[5] M. Shoush, M. Dumas, Prescriptive process monitoring under resource constraints: A reinforcement learning approach, arXiv (2024).

[6] I. Teinemaa, N. Tax, M. de Leoni, M. Dumas, F. M. Maggi, Alarm-based prescriptive process monitoring, in: M. Weske, M. Montali, I. Weber, J. vom Brocke (Eds.), Business Process Management Forum, Springer International Publishing, Cham, 2018, pp. 91–107.

[7] S. Branchi, C. Di Francescomarino, C. Ghidini, D. Massimo, F. Ricci, M. Ronzani, Learning to act: A reinforcement learning approach to recommend the best next activities, in: C. Di Ciccio, R. Dijkman, A. del Río Ortega, S. Rinderle-Ma (Eds.), Business Process Management Forum, Springer International Publishing, Cham, 2022, pp. 137–154.

[8] M. S. Qafari, W. van der Aalst, Root cause analysis in process mining using structural equation models, in: A. Del Río Ortega, H. Leopold, F. M. Santoro (Eds.), Business Process Management Workshops, Springer International Publishing, Cham, 2020, pp. 155–167.

[9] S. J. J. Leemans, N. Tax, Causal reasoning over control-flow decisions in process models, in: X. Franch, G. Poels, F. Gailly, M. Snoeck (Eds.), Advanced Information Systems Engineering, Springer International Publishing, Cham, 2022, pp. 183–200.

[10] C. Gong, D. Yao, C. Zhang, W. Li, J. Bi, L. Du, J. Wang, Causal discovery from temporal data, in: Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, KDD '23, Association for Computing Machinery, New York, NY, USA, 2023, p. 5803–5804. URL: https://doi.org/10.1145/3580305.3599552. doi:10.1145/3580305.3599552.

[11] J. De Smedt, F. Hasić, S. K. vanden Broucke, J. Vanthienen, Holistic discovery of decision models from process execution data, Knowledge-Based Systems (2019). doi:https://doi.org/10.1016/j.knosys.2019.104866.

[12] O. López-Pintado, M. Dumas, J. Berx, Discovery, simulation, and optimization of business processes with differentiated resources, Information Systems 120 (2024) 102289. URL: https://www.sciencedirect.com/science/article/pii/S0306437923001254. doi:https://doi.org/10.1016/j.is.2023.102289.

[13] A. Maldonado, C. M. M. Frey, G. M. Tavares, N. Rehwald, T. Seidl, Gedi: Generating event data with intentional features for benchmarking process mining, in: A. Marrella, M. Resinas, M. Jans, M. Rosemann (Eds.), Business Process Management, Springer Nature Switzerland, Cham, 2024, pp. 221–237.

[14] M. Pourbafrani, S. Vasudevan, F. Zafar, Y. Xingran, R. Singh, W. M. P. van der Aalst, A python extension to simulate petri nets in process mining, 2021. arXiv:2102.08774.

[15] Z. D. Bozorgi, I. Teinemaa, M. Dumas, M. La Rosa, A. Polyvyanyy, Process mining meets causal machine learning: Discovering causal rules from event logs, in: 2020 2nd International Conference on Process Mining (ICPM), 2020, pp. 129–136. doi:10.1109/ICPM49681.2020.00028.

[16] I. Bica, A. M. Alaa, J. Jordon, M. van der Schaar, Estimating counterfactual treatment outcomes over time through adversarially balanced representations, in: International Conference on Learning Representations, 2020.

[17] W. M. P. van der Aalst, Object-centric process mining: Unraveling the fabric of real processes, Mathematics (2023).
[18] V. Mai, K. Mani, L. Paull, Sample efficient deep reinforcement learning via uncertainty estimation, arXiv preprint arXiv:2201.01666 (2022).
[19] S. Levine, A. Kumar, G. Tucker, J. Fu, Offline reinforcement learning: Tutorial, review, and perspectives on open problems, arXiv (2020).