Towards Organizational Digital Twins: Process Enhancement by Object-Centric Simulation

Benedikt Knopp

Chair of Process and Data Science, RWTH Aachen University, Ahornstraße 55, 52074 Aachen, Germany

Abstract

Simulation is used for the enhancement of businesses by anticipating the potential course of business processes and thus supporting strategical and operational decision-making. Existing methods assume a fixed case notion of processes, and thus fail to capture the nature of real organizational processes as an interplay of objects of various types. Through object-centric approaches in process mining, it is possible to overcome this limitation. This Ph.D. thesis aims at developing simulation methods on top of object-centric process models, providing means for decision support of complex processes, as well as enabling the implementation of simulation-based Organizational Digital Twins (DTOs) for automated process enhancement.

Keywords

process enhancement, process simulation, object-centricity, digital twins

1. Introduction

Companies use simulation to anticipate the impact of changing environmental or company-internal parameters, for example, concerning supply and demand or resource availability, on their processes. By this, decisions for improvement actions in operations are supported. Through advances in process mining, it is possible to create models for discrete event simulation that are highly reliable by exploiting historic event data. A limitation of these models is that they consider single processes in isolation and fail to capture the nature of real business processes as a composite of nested and concurrent smaller processes. For example, a classical order-to-cash macro-process consists of subroutines in accounting, production and logistics. The paradigm of object-centricity [1] in process mining aims to remedy this limitation by means of a holistic modeling and analysis of processes as an interplay of objects of various types. The first goal of this thesis is to enrich such object-centric models to obtain executable *object-centric process simulation (OCim)* models that provide means for a more holistic decision support.

The second goal of this thesis is to contribute to automated process enhancement as follows. The envisioned simulation methods can be seen as an enabling technology for *Digital Twins (DTs) of organizations (DTOs)* [2]. A DT is a monitor of a system that autonomously runs simulations to foresee the effects of changing system parameters, while also linking back to the system by automatically triggering improvement actions based on the simulation results. A DTO scopes

ICPM 2023 Doctoral Consortium and Tool Demonstration Track
knopp@pads.rwth-aachen.de (B. Knopp)
0000-0002-7762-9486 (B. Knopp)

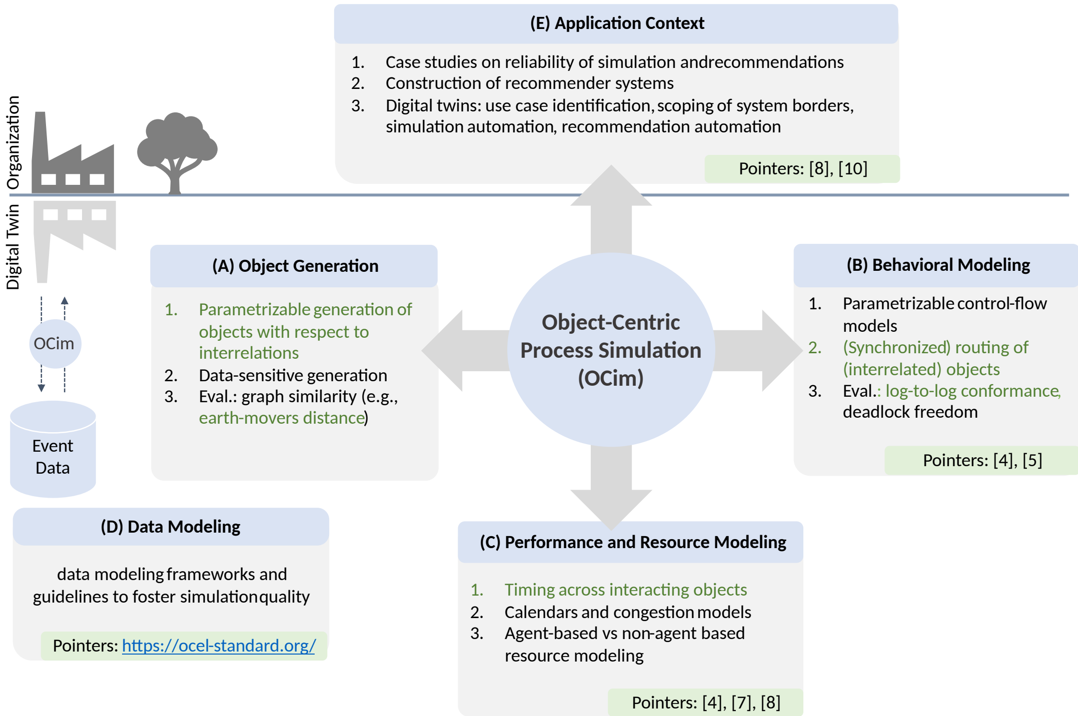

Figure 1: Categories of action points for (automated) process enhancement using object-centric process simulation. Action points marked green have been approached in our proof-of-concept [3]. We also name pointers that will steer future work in the respective direction.

over an organization to the extent that is necessary to capture all relevant information of a macro-process, such as order-to-cash. Thus, OCim provides a missing link for DTOs by enabling both the holistic scoping through object-centricity, and the improvement capabilities through simulation.

2. Research Goals

The goal of this thesis is to develop methods for executing OCim and for evaluating these methods concerning the reliability of simulation runs and their applicability in business contexts. Towards these goals, we identify the work packages *(A)* to *(E)* outlined in Fig. 1. *(A), (B), (C)* address the components of object-centric simulation models, namely *(A)* a case generator and *(B), (C)* the simulation runtime. *(D)* is concerned with research on the underlying model for object-centric event data. *(E)* is to put the developed methodology into application context.

In the remainder of this abstract, the work packages *(A) to (E)* are specified. Firstly, the maturity of a baseline solution for OCim is discussed (Sec. 3), in order to mark completed action points. Open or incompletely solved action points are discussed in Sec. 4. Finally, our proposed research is compared against other work (Sec. 5).

3. Completed Research

In [3], we present a method for discovering OCim models from event data. Artificial objects are generated *(A1)*, resembling the nature of observed objects and their interrelations. For instance, in an order-to-cash process, the number of items (first object type) per customer order (second object type) is an object interrelation. The novelty of our approach is given because these interrelations are made parametrizable; for instance, by increasing the average number of items per order. Our approach then is to route the generated objects synchronously through a control-flow model with respect to their interrelations *(B2)*. An added value of our approach is that this simulation allows to examine how an adapted parametrization affects the behavior or performance of individual object types, for instance, the effect on cycle times of orders if customers order more items *(A1, C1)*. We empirically showed the validity of our approach by measuring conformance between flattened input and output logs *(B3)*, similarity between sets of objects *(A3)* as well as average cycle times.

4. Open Challenges

Regarding the solutions for *(A)*, *(B)* and *(C)*, we observe shortcomings in [3] and open action points for these as well as work packages *(C)*, *(D)* as follows.

4.1. OCim Framework (A, B, C)

For case generation *(A)*, [3] provides an iterative method that makes object multiplicities parameterizable. First, this will be extended to include object attributes *(A2)*. A requirement here is that object generation should also be sensitive to these attributes; for instance, orders are likely to comprise more items if these items are cheaper. Second, other methods (such as sampling) for object generation will be tested *(A1)*.

Concerning the behavioral component *(B)*, [3] uses simple stochastic methods based on flattened event logs for controlling the routing of objects. To refine this, we will leverage designated feature representations for object-centric (graph-based) data [4]. More refined predictive models could generalize Data Petri nets enriched with decision trees, as proposed in [5], to the object-centric nets. This will also enable a structured parametrization of the simulation by manipulating decision rules *(B1)*. One challenge here is to develop joint decision rules, i.e. , decisions involving multiple objects. The use of object-centric features may simultaneously serve for a more realistic representation of timing behavior *(C1)*. Concerning evalation, we need to assess whether the configuration of a simulation model allows for a deadlock-free execution [6], and develop designated log-to-log conformance metrics for the object-centric setting.

Concerning the performance model *(C)*, [3] models execution delays of activities based on flattened object types, also evaluating timing aspects per object type *(C1)*. While object type wise evaluation of performance is already useful, the modeling of performance and also of resources clearly needs to respect object interactions. For instance, the service time of packing a handling unit depends on the number of items to be packed. It will also be interesting to explore the modeling of resources as object types and create, for example, object-centric congestion models [7] *(C2)*. Work on resource modeling, however, needs to be sensitized for the difficulties of modeling agent-based systems *(C3)* [8].

4.2. Data Model (D)

The quality of data-driven simulation highly depends on the quality and the structure of the underlying data *(E)*. In [3], we relied on a standard [9] that does not explicitly represent object relations. Therefore, object interaction patterns are implicit through event co-occurrence. Thus, we will leverage explicit object relations based on advanced data modeling[1]. Also, based on case studies, we hope that we can formulate modeling guidelines for event data that facilitate data-driven simulation or object-centric process analysis in general.

4.3. Application Context (E)

Finally, our methods are to be utilized in the scope of a DTO. A DTO as introduced in Sec. 1 encompasses automated improvement capabilities. This implies the need for a bi-directional control-flow between model and system. As a first step towards this challenging goal, we will design recommender systems *(E1)*, because efforts towards automatized implementation of corrective actions should be preceded by work on supplying these corrective actions as recommendations to a human-in-the-loop. To this end, we will extend case studies on the technical quality of our framework with a study on the quality of recommendations [10] *(E2)*.

The creation of a DTO *(E3)* requires (a) a concrete use case and business goal, (b) an according scoping the borders of the system of interest and (c) the automation capabilities of simulation and recommendation as described above. To this end, we propose a modular and incremental approach: these challenges will be tackled first independently to be then integrated into a digital twin. An incremental approach also applies to the subgoal of scoping the system borders. Here, we will consider to create models of micro-processes of organizations first and extend these towards macro-models.

5. Related Work

The basic idea of data-driven process simulation, namely automated simulation model discovery, is best described in the fundamental work of Rozinat et al. [11]. Meanwhile, such solutions are well-matured in the classical non-object centric setting [12, 13], also providing capabilities for parameter optimization. Note that this thesis aims at generalizing such methods. Therefore, the methods developed herein should also be validated by checking whether they generalize the capabilities of existing simulation tools and metrics [14].

Acknowledgments

Funded under the Excellence Strategy of the Federal Government and the Länder. I thank the Alexander von Humboldt (AvH) Stiftung for supporting our research.

[1]https://ocel-standard.org/

References

[1] 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.

[2] R. Parmar, A. Leiponen, L. D. Thomas, Building an organizational digital twin, Business Horizons 63 (2020) 725–736.

[3] B. Knopp, M. Pourbafrani, W. M. P. van der Aalst, Discovering object-centric process simulation models, in: International Conference on Process Mining (ICPM), Springer, 2023.

[4] 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.

[5] M. de Leoni, W. M. P. van der Aalst, Data-aware process mining: discovering decisions in processes using alignments, in: Proceedings of the 28th annual ACM symposium on applied computing, 2013, pp. 1454–1461.

[6] K. Batoulis, S. Haarmann, M. Weske, Various notions of soundness for decision-aware business processes, in: Conceptual Modeling: 36th International Conference, ER 2017, Valencia, Spain, November 6–9, 2017, Proceedings 36, Springer, 2017, pp. 403–418.

[7] A. Senderovich, J. C. Beck, A. Gal, M. Weidlich, Congestion graphs for automated time predictions, in: Proceedings of the AAAI Conference on Artificial Intelligence, volume 33, 2019, pp. 4854–4861.

[8] M. C. Becker, B. T. Pentland, Digital twin of an organization: Are you serious?, in: International Conference on Business Process Management, Springer, 2021, pp. 243–254.

[9] 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.

[10] M. Dees, M. de Leoni, W. M. P. van der Aalst, H. A. Reijers, What if process predictions are not followed by good recommendations?(technical report), arXiv preprint arXiv:1905.10173 (2019).

[11] A. Rozinat, R. S. Mans, M. Song, W. M. P. van der Aalst, Discovering simulation models, Information systems 34 (2009) 305–327.

[12] M. Camargo, M. Dumas, O. González-Rojas, Automated discovery of business process simulation models from event logs, Decision Support Systems 134 (2020) 113284.

[13] M. Pourbafrani, S. Jiao, W. M. P. van der Aalst, Simpt: process improvement using interactive simulation of time-aware process trees, in: International Conference on Research Challenges in Information Science, Springer, 2021, pp. 588–594.

[14] D. Chapela-Campa, I. Benchekroun, O. Baron, M. Dumas, D. Krass, A. Senderovich, Can I trust my simulation model? measuring the quality of business process simulation models, in: Business Process Management - 21st International Conference, BPM 2023, volume 14159, Springer, 2023, pp. 20–37.