Integration of Process Mining and Simulation: A Survey of Applications and Current Research

Brian Keith Norambuena(✉)

Departamento de Ingeniería de Sistemas y Computación,
Universidad Católica del Norte, Av. Angamos 0610, Antofagasta, Chile
brian.keith@ucn.cl

Abstract. Process mining is a field that uses elements from data mining and business process modeling to perform tasks such as process discovery, conformance checking, and process enhancement. This article presents a study about the application of process mining techniques in simulation and the integration between the two disciplines. Specifically, it shows a series of developments in the field that illustrate the possible applications. Also, the current main challenges of integrating process mining and simulation are exposed, one of the main issues being the lack of compatibility between tools for process mining and simulation. The objective of this article is to show the importance and practical utility of applying process mining approaches in simulation. The literature review shows that while there have been developments towards an integrated framework for process mining and simulation, the development of a standard unified methodology is still an open problem.

Keywords: Process mining · Simulation · Process improvement

1 Introduction

Process mining is a discipline that lies between data mining and process modelling. Its main objective is to discover processes, perform conformance checking and process improvement. Process mining seeks automatizing these three tasks by applying data mining techniques specially designed for dealing with process data [1, 2]. The main tasks of process mining are summarized in Fig. 1.

Analyzing a process automatically and determining where deviations are produced may be used for taking corrective actions on how processes are done. This corresponds to the process improvement phase, where corresponding improvements are proposed, as a function of the information obtained previously [1].

Computer simulation in general encompasses different challenges, in the context of process mining, this work will center on *business process simulation*. This field interprets a business as a series of inter-related processes and works with the underlying assumption that these processes are composed of activities that transform inputs into outputs. The outputs from each process can be analyzed in order to look for opportunities of process enhancement [3].

Á. Rocha et al. (Eds.): WorldCIST'18 2018, AISC 745, pp. 287–294, 2018.
https://doi.org/10.1007/978-3-319-77703-0_29

Although both fields are related, the key difference between these disciplines is that while process mining allows the user to understand the current or past behavior of a process, simulation is intended to predict future behavior. However, understanding a process and predicting its outcomes are, of course, closely related. Considering this, these two disciplines can be seen as complementary to each other and as such they can be used in conjunction to improve processes in organizations. In a sense, this gap is similar to the one found between building analytical models and data-driven models in other fields.

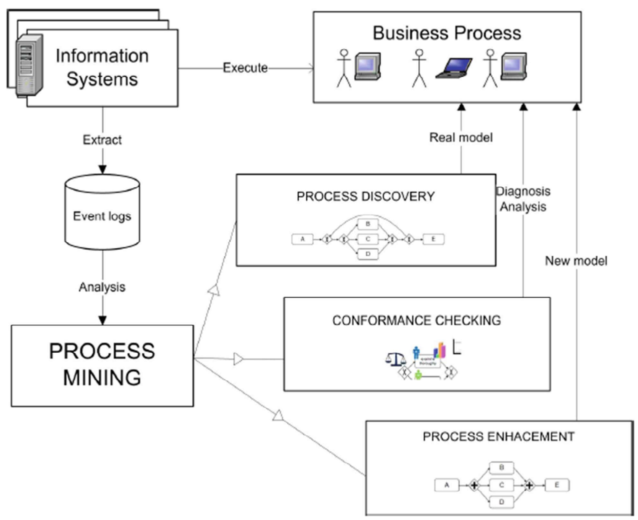

Fig. 1. The main tasks of process mining (source: reproduced from [1]).

The Process Mining Manifesto states that one of the major challenges that need to be addressed is the usability of process mining techniques and its integration with other process improvement methods [2, 22]. Integration with simulation tools and techniques are considered part of this challenge. The objective of this article is to explore recent developments in the integration of process mining and simulation techniques, this is done through a systematic review of the literature.

The remaining parts of this document are structured as follows. First, a section on process mining and simulation describes the reasons that lead to the natural requirement of integration. Then, selected case studies and research from the literature are described in order to illustrate the integration. After this, the challenges associated with integrating simulation and process mining are discussed with the perspective gained from the case studies. In the final section, the conclusions from this research are presented.

2 Process Mining and Simulation

2.1 The Pitfalls of Simulation

The idea of applying process mining to complement simulation techniques arises naturally when the recurrent pitfalls of simulation are analyzed. In accordance to the literature, a list of common flaws in contemporary simulation approaches is presented. These problems arise from both the human elements during the process modeling or due to intrinsic methodological problems [4, 5], a summary of these is provided in Table 1.

Table 1. Summary of the common pitfalls in simulation modeling.

Problem	Description	Source
Interviews	Interviews with business experts can result in contradictory information, also interviewees perception tends to be biased to a certain extent	Human-centered
Hawthorne effect	When using observational data, the Hawthorne effect can occur (i.e. changes in behavior caused by the presence of observers)	Human-centered
Modeling strategy	Modeling from scratch rather than using existing artifacts, disregarding the information present in these elements leads to mistakes and unnecessary work. It must also be noted that process documentation may deviate from the real process behavior	Methodological
Modeling focus	Focus on design rather than operational decision making, which is helpful for the initial design of a business process but less suitable for operation decision making and continuous improvement	Methodological
Modeling details	Insufficient modelling of resources: the behavior of resources is typically modelled in a rather naïve manner	Methodological

2.2 Integration with Process Mining

The high-level idea of process mining integration with simulation is described in Fig. 2. In particular, the event logs are used to obtain the AS-IS process model through process mining. After this several new alternative processes can be designed (TO-BE process models). Using simulation both the original model and the new alternatives can be evaluated, through a comparative analysis a new improved process model is selected, leading to its implementation in the organization.

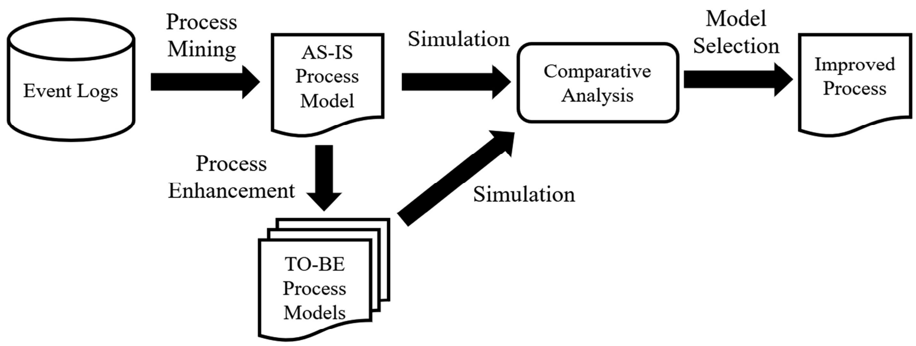

Fig. 2. High level idea behind the integration of process mining and simulation (source: own elaboration).

While other integration strategies may be applied for simulation and process mining, this form arises naturally and addresses the main concerns shown in the previous subsection, particularly the human-related activities. However, it must be noted that to implement this, it is necessary to have event logs available or data that follows the process mining meta-model [6] shown in Fig. 3.

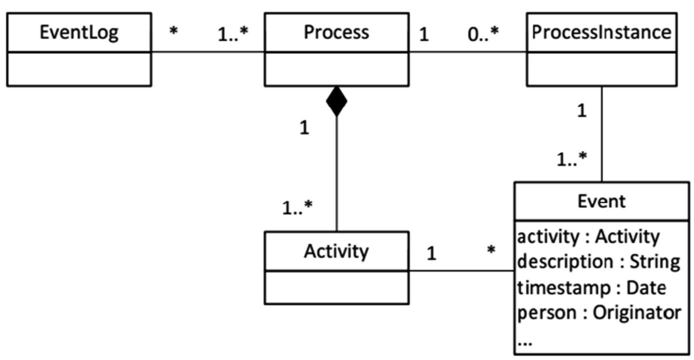

Fig. 3. Process mining metamodel (source: reproduced from [6]).

In Fig. 3 each process consists of activities and is associated with a series of process instances. Each process instance consists of one or more events. These events must indicate the activity they correspond to and have a description, a time to be executed, and a person in charge. If this data is not available, extracting the AS-IS process models may not be feasible.

3 Applications

There exist several attempts at creating a methodology that unifies process mining and simulation. Yet the goal of finding a standard methodology to do this remains elusive. This section shows the results of some studies found in the literature. These studies were selected according to their illustrative value. Only the most important results from these papers are presented here, an in-depth account of the results can be found in the referenced works.

Rozinat et al. approach the challenge of discovering simulation models using process mining techniques in a semi-automatic fashion. Their work demonstrates that it is possible to create simulation models based on event logs that encompass different perspective (e.g. control-flow, data, resources or time). Their work has been implemented through ProM [7] and has been evaluated thoroughly [8].

In [9], Liu explores the integration of process mining and discrete event simulation. As stated by Liu process mining tools can easily be used to characterize operational processes through the use of historical data. However, these tools do not provide an interpretable way to import the results simulation tools in an easy way. On the other hand, simulation models have a hard time replicating existing processes. Liu has taken a first step towards bridging the gap between process mining and discrete event simulation, by enabling the compatibility between the tools of both disciplines [9].

The use of process mining and simulation in conjunction has been studied in the context of process redesign projects for service operations [10]. These operations require custom-tailored analysis techniques for the modeling, because they present unique characteristics (e.g. balance between quality-of-service and resource efficiency). The author of [10] presents a methodological framework using queuing theory, since it fits in a natural way with the characteristics of service operations. Their main contribution is the queuing perspective for mining simulation processes.

In [11] process mining is applied with an agent-based simulation [12, 13] approach to business processes. Agent-based simulation is focused on the analysis of organizational systems through the use of agents that communicate and interact with each other. These agents represent business actors in the processes. This work applies the AOR simulation framework [14] for its different abstraction levels (e.g. activities and message events can both be defined) and it shows that it is possible to apply agent-based simulation for the execution of business processes and consequent log generation. An iterative method of process improvement is proposed based on process mining and agent-based simulation. This integration arises from the different abstraction levels of process mining and agent-based models [11].

Process models created by humans can be an idealistic version of the process and can suffer from several flaws, due to the complex nature and human element of process modeling. In this context, the analysis of an information system's event log (through process mining) can be useful to help in the modeling step. Using simulation techniques and both process and data mining methods in conjunction, the authors of [15] propose a methodological approach to process redesign based on the work of Măruşter [16] and complementing the methodology described therein. In this context, process mining provides the required parameters for the creation of the simulation model, based upon a realistic model of the process extracted from the event logs. Then, a comparison of different redesign alternatives can be achieved through simulation [15].

Finally, workflow environments used for process mining can be complemented with simulation tools [17]. Adding these capabilities allows for predictive modeling, activity monitoring and real-time optimizations [18]. In general, workflow based environments provide an advantageous approach to process mining and simulation [17, 21].

4 Challenges

While there have been several developments in the integration of process mining and simulation, there are several extant issues that must be addressed before a widespread adoption of these techniques can be seen in industry. The first and foremost challenge is of course the lack of compatibility between the different tools used for process mining and other fields [9]. However, there are other general challenges faced by process mining [1, 2, 15, 19, 20], these challenges are presented and described in Tables 2 and 3.

Table 2. Process mining challenges from an organizational and data-centric perspective.

Problem	Description
Process knowledge	Information systems may be agnostic to business processes. This means that actions performed in the systems are kept in some record, but the systems do not know to which process these activities correspond to. In the case of process-agnostic systems, their records may not have enough data to apply process mining algorithms. [1, 2, 20]
Unstructured data	Process mining strongly depends on the quality of data collected and stored. Big amounts of data must be frequently filtered due to incomplete process instances. Process mining efficiency and effectiveness can be substantially benefited from well-structured and well-defined datasets and also collection guides that allow obtaining high-quality datasets [19]
Noisy data	Another problem in datasets is noise. This may sometimes arise due to unexpected processes (activities which are not supposed to occur or occurring in incorrect positions) [20]
Abstraction level	Due to the characteristics of the data or the source information systems, in some cases the amount of details is excessive, and the complexity of the resulting model makes it impractical. The opposite also holds, in the sense that there may not be enough data to provide a detailed model. In general, every good model must keep a balance between precision and the representation of reality and computational complexity [2, 20]

Most challenges can be classified as either methodological or data-centric. Data-centric challenges may be addressed through the standardization of the data generated by information systems and an adequate system design. Methodological challenges require, as their name indicates, the development of new methodoligies (or improvement of existing ones whenever possible).

Table 3. Process mining challenges from a methodological perspective.

Problem	Description
Result evaluation	The assessment of the process obtained through process mining techniques is a challenge, because it is necessary to define a standard and rigorous procedure to assess its quality [1]
Optimization	Another challenge for the implementation of process mining techniques is that algorithms must be properly parametrized. The search of optimum parametrization can be a complex task due to the number of available parameters [1, 20]
Usability	As mentioned in this paper and in the corresponding literature, integration with other methodologies and analytical techniques is still an open issue [1, 2, 15]

In spite of all these challenges, advances have been made towards the integration of process mining with simulation. However, a complete framework or methodological treatise still remains an important issue to be resolved.

5 Conclusions

This study addresses the main aspects necessary to understand the importance of process mining and its integration with process simulation. In addition, the different challenges and problems faced by process mining are identified. On the other hand, the study shows that process mining may be applied both for building the simulation model as well as for validation of a hand-crafted model. Several applications illustrating these concepts are shown.

Since process mining is a relatively new field compared with the disciplines forming it, its growth potential is quite high, both on a theoretical and practical basis, and as such, it is expected that better tools and integration capabilities will be developed in the coming years. The challenges that must be faced in the field of process mining are also described. One of the main challenges is the availability of suitable data provided by information systems aware of business processes.

Although there are methodologies to apply process mining techniques in general, the development of a standard methodology to integrate process mining techniques with simulation in organizations is still an open problem. Finally, the importance of the integration between simulation and process mining is discussed.

References

1. Van Der Aalst, W.: Process Mining: Discovery, Conformance and Enhancement of Business Processes. Springer Science & Business Media, Heidelberg (2011)
2. Van Der Aalst, W., Adriansyah, A., de Medeiros, A.K.A., Arcieri, F., Baier, T., Blickle, T., Bose, J.C., van den Brand, P., Brandtjen, R., Buijs, J., et al.: Process mining manifesto. In: Business Process Management Workshops, pp. 169–194. Springer (2011)
3. Tumay, K.: Business process simulation. In: Proceedings of the 28th Conference on Winter Simulation, pp. 93–98. IEEE Computer Society (1996)

4. Nakatumba, J., Rozinat, A., Russell, N.: Business process simulation: how to get it right. In: International Handbook on Business Process Management. Citeseer (2009)
5. Martin, N., Depaire, B., Caris, A.: The use of process mining in a businesss process simulation context: overview and challenges. In: 2014 IEEE Symposium on Computational Intelligence and Data Mining (CIDM), pp. 381–388. IEEE (2014)
6. Poncin, W., Serebrenik, A., Van Den Brand, M.: Process mining software repositories. In: 2011 15th European Conference on Software Maintenance and Reengineering (CSMR), pp. 5–14. IEEE (2011)
7. Van Dongen, B.F., de Medeiros, A.K.A., Verbeek, H., Weijters, A., Van Der Aalst, W.M.: The ProM framework: a new era in process mining tool support. In: ICATPN, vol. 3536, pp. 444–454. Springer (2005)
8. Rozinat, A., Mans, R.S., Song, M., van der Aalst, W.M.: Discovering simulation models. Inf. Syst. **34**(3), 305–327 (2009)
9. Liu, S.: Integrating Process Mining with Discrete-Event Simulation Modeling. Brigham Young University, Provo (2015)
10. Senderovich, A.: Service analysis and simulation in process mining. In: International Conference on Business Process Management, pp. 578–581. Springer (2014)
11. Szimanski, F., Ralha, C.G., Wagner, G., Ferreira, D.R.: Improving business process models with agent-based simulation and process mining. In: Enterprise, Business-Process and Information Systems Modeling, pp. 124–138. Springer (2013)
12. Bonabeau, E.: Agent-based modeling: methods and techniques for simulating human systems. Proc. Natl. Acad. Sci. **99**(suppl 3), 7280–7287 (2002)
13. Baqueiro, O., Wang, Y.J., McBurney, P., Coenen, F.: Integrating data mining and agent based modeling and simulation. In: Industrial Conference on Data Mining, pp. 220–231. Springer (2009)
14. Wagner, G.: AOR modelling and simulation: towards a general architecture for agent-based discrete event simulation. In: Agent-Oriented Information Systems, pp. 174–188. Springer (2004)
15. Aguirre, S., Parra, C., Alvarado, J.: Combination of process mining and simulation techniques for business process redesign: a methodological approach. In: Data-Driven Process Discovery and Analysis, pp. 24–43. Springer (2012)
16. Măruşter, L., van Beest, N.R.: Redesigning business processes: a methodology based on simulation and process mining techniques. Knowl. Inf. Syst. **21**(3), 267 (2009)
17. Wynn, M., Rozinat, A., van der Aalst, W., ter Hofstede, A., Fidge, C.: Process mining and simulation. In: Modern Business Process Automation, pp. 437–457. Springer (2010)
18. Aguilar, M., Pater, A.J.G.: Business process simulation: a fundamental step supporting process centered management. In: Farrington, P., Nembhard, H., Sturrock, D., Evans, G. (eds.) Proceedings of the 1999 Winter Simulation Conference (WSC 1999), Phoenix, AZ, USA, December 1999, pp. 1383–1392. ACM (1999)
19. Bose, R., Mans, R.S., van der Aalst, W.M.: Wanna improve process mining results? In: 2013 IEEE Symposium on Computational Intelligence and Data Mining (CIDM), pp. 127–134. IEEE (2013)
20. Burattin, A.: Obstacles to applying process mining in practice. In: Process Mining Techniques in Business Environments, pp. 59–63. Springer (2015)
21. Rozinat, A., Wynn, M.T., van der Aalst, W.M., ter Hofstede, A.H., Fidge, C.J.: Workflow simulation for operational decision support. Data Knowl. Eng. **68**(9), 834–850 (2009)
22. Gonçalves, J., Faria, B.M., Reis, L.P., Carvalho, V., Rocha, Á.: Data mining and electronic devices applied to quality of life related to health data. In: 2015 10th Iberian Conference on Information Systems and Technologies (CISTI), pp. 1–4. IEEE (2015)