Constructing Digital Twins for Accurate and Reliable What-If Business Process Analysis

Marlon Dumas

University of Tartu, Tartu, Estonia
marlon.dumas@ut.ee

Abstract. A long-standing problem in the field of Business Process Management (BPM) is that of constructing accurate and reliable models for "what-if" business process analysis (digital process twins). This paper formulates this problem in a general setting and spells out associated challenges. The paper suggests that this problem can be addressed by combining observational data, experimental data, and domain knowledge using hybrid modeling methods drawing from the fields of discrete event simulation, machine learning, and causal inference.

1 The Problem

Business processes are destined to change, be it to respond to business opportunities, environmental changes, changes in customer expectations, staff turnover, or to address internal bottlenecks and other sources of waste [9].

Changes in a process may be *organic* — resulting from actors in a process adjusting their behavior to respond to emergent situations — or *planned* — resulting from agreed-upon changes to the norms, guidelines, policies, or IT systems. In this paper, we focus planned changes [9], herein called *interventions*. Examples of interventions include reordering two tasks, adding a task, adding a resource, changing the decision logic in a branching point, or automating a task.

Interventions are usually made in order to improve the performance of a process or to comply with regulations. In either case, process managers are interested in predicting and understanding the impact of the changes with respect to one or more process performance measures, such as *cycle time* (duration of each case averaged across a set of cases), *activity processing time* (total time consumed by an activity type across a set of cases) and *resource utilization* (percentage of available time during which a resource is busy).

A longstanding problem in the field of BPM is that of *what-if process analysis*: How to predict the values that one or more process performance measures will take after a given business process intervention?

An intervention on a process may effect other processes, particularly those with which it shares resources. For example, the same team of accountants might handle inbound invoices in a Purchase-to-Pay process (P2P) process and outbound invoices in an Order-to-Cash (O2C) process. Thus, an intervention in the P2P process that increases the team's workload may affect the O2C process.

Hence, rather than viewing an intervention as affecting an individual process, we should reason about one or more interventions affecting one or more processes. In the above example, the intervention in the P2P process should be analyzed relative both to the P2P and the O2C process. Accordingly, we formulate the problem of what-if process analysis as follows:

- Given:
  1. one or more business processes, for which we might have available one or more process specifications and/or event logs generated by the execution of the processes on top of one or more information systems.
  2. Given one or more interventions that affect one or more processes.
  3. Given one or more process performance measures of interest.
- Predict the values of the process performance measures after the given interventions.

A solution to this problem should fulfill the following requirements:

R1 The predictions about the effects of interventions should be accurate. Here, accuracy may be measured in terms of the error between the predicted and the actual performance after intervention.

R2 Sometimes, it might not be possible to predict the effect of a given intervention, for example because no similar interventions have been observed in the past, and thus there is no information about how the actors in the process will react. If this is the case, either no prediction should be made or the prediction should be accompanied by a reliability estimate.

From a design science perspective, the problem at hand can be formulated as that of designing a method to construct an artifact that can make the above predictions for a given set of processes, possible interventions, and performance measures. Said method may take as input data and domain knowledge about one or more processes, for example, execution data (event logs), constraints on the process (declarative specifications), or other specifications of process behavior.

The above design science problem is not specific to the field of BPM. A similar problem has been tackled for several decades in various fields of engineering, including mechanical and industrial engineering, where the type of artifact mentioned above is known under the term of *Digital Twin* (DT). A digital twin is a model of an object or system that, together with a stream of data about events related to the object or system, is able to accurately predict the performance of the physical object or system over time. Hence, the problem we formulate here can be conceptualized as that of designing methods to construct a DT of a business process, herein called a Digital Process Twin (DPT).

2 Why is the problem important?

During the redesign phase of the business process lifecycle, what-if analysis allows decision makers to compare multiple possible interventions in order to address one or more performance issues in a process [6]. What-if analysis allows managers to estimate the performance improvements that a given process intervention is likely to bring about, and hence to build a business case for an intervention by comparing its expected benefits against its expected costs or drawbacks.

In addition to its role as a decision making during process redesign, what-if analysis is a primitive operation for automated business process optimization. Here automated process optimization refers to the ability to identify optimal interventions on a process to maximize or minimize a given objective function (defined in terms of one or multiple performance measures) under certain constraints (e.g. resource utilization should remain below 80%)

DPTs also have potential applications in the field of process optimization [10]. Process optimization algorithms operate by exploring a large number of possible process interventions and evaluating them in order to determine which intervention (or combination of interventions) yields the highest gain with respect to the given objective function. To do so, the optimization algorithm needs a DPT that is sufficiently efficient (computationally) to be executed thousands or millions of times in a short period of time. This use-case brings an additional requirement in the above problem statement: the DPT should be computationally efficient, particularly when it comes to generating predictions.

3 Why is the problem challenging?

The field of DT construction has reached a certain level of maturity in various engineering fields (e.g. mechanical and material engineering) where the objects for which DTs are built are guided by physical laws. There have also been considerable advances in the field of DT construction for deterministic systems such as automated production systems, where DTs are often built using Discrete Event Simulation (DES) [12]. However, the more a system involves a human factor — and business processes do involve a significant human component — the more intractable it becomes to design digital twins of sufficiently high level of fidelity. The underpinning challenge is twofold.

First, human behavior exhibits a high level of variability. The performance of a worker on 1 June might be very different from that on 8 June, at exactly the same time. Some of these variations can be predicted if suitable data is available: For example, the time since the worker's last vacation period, the weather, the level of workload, the fact that 8 June is a school holiday. Others, however, are stochastic, e.g. fatigue effects due to a viral infection or a bad night's sleep. Beyond individual variability (variability in the behavior of the same worker), there is also cross-individual variability: different workers behave differently under indistinguishable circumstances. A more experienced worker is likely to complete a task instance faster than a less experienced one, under identical process execution contexts. There are two intertwined challenges when modeling such variability: (i) separating the predictable variability from the unpredictable (stochastic) one; and (ii) collecting the data required to model the predictable component, e.g. collecting suitable data to train Machine Learning (ML) models to capture fatigue-related or workload-related variability.

Surmounting the variability challenge is necessary to build simulation models that accurately capture an "as-is" process. However, the purpose of a DPT is not so much to capture the as-is process, but rather to reason about the effects of interventions, including interventions that have not been observed before. Yet, human actors may react to interventions in unexpected ways [5]. Automating a task might slow down workers downstream if this automation generates issues that would otherwise be detected by the workers who performed said task before its automation. Hence, a second challenge is how to reliably and accurately model the way the performance of the process evolves after an intervention.

4 Related work and research directions

Business Process (BP) simulation is an established approach for what-if analysis. Traditional BP simulation methods have various limitations that affect their accuracy [14]. For example, BP simulators often assume that resources are available during predetermined timetables, that there is no multi-tasking, no batching, and no task prioritization. Beyond these limitations, BP simulation approaches suffer from two fundamental assumptions: (i) they assume robotic resource behavior (e.g. do not take into account fatigue); and (ii) they rely on manually designed models that often only capture the main pathways of the process. Data-driven simulation approaches [3] — wherein the simulation model is automatically discovered from execution data — address some of these shortcomings, e.g. in relation to handling resource availability patterns and multitasking [2]. A direction to enhance the accuracy of data-driven BP simulation methods is to combine them with methods for automated discovery of batching [7], prioritization [11], and other behaviors that affect the accuracy of simulations [14].

A related approach to build DPTs is via system dynamics simulation [8]. Whereas BP simulations capture the behavior of each case, these approaches focus on relations between measures (e.g. resource utilization, cycle time).

Another possible approach to build DPTs is via ML methods. Recent studies have shown that it is possible to build generative machine learning models that accurately replicate the behavior of a process under different perspectives, particularly the control-flow and the temporal perspective [13]. Recent work showed that such generative models can achieve higher level of temporal accuracy than BP simulation models derived from execution data [4]. However, generative models are trained on past data, and are therefore unable to make predictions for interventions that have not been observed in the past. In contrast, DES models are able to capture the effect of certain types of interventions (particularly interventions that affect the availability of resources). A direction for building accurate DPTs is to combine generative ML approaches with existing (data-driven) BP simulation approaches. The challenge here is how to do so in a way that the resulting models have the generalizability required for what-if analysis, beyond simple changes such as removing a task, replacing a task with another one, or adding a task that ressembles existing tasks in terms of its behavior.

Another direction is to combine data-driven simulation with causal ML techniques. Recent work showed that causal ML techniques (uplift trees) can be used to assess the impact of interventions on case outcomes [1]. Similar methods can be used to estimate the impact of interventions on cycle time. An advantage of these techniques is that their reliability can be enhanced by combining observational data (event logs) with experimental (A/B testing) data. The challenge is how to seamlessly integrate these techniques with simulation in order to leverage their complementary, given that causal ML techniques can only reason about interventions that have been observed, in contrast to simulation.

Acknowledgment Work funded by European Research Council (PIX Project).

References

1. Z. D. Bozorgi, I. Teinemaa, M. Dumas, M. L. Rosa, and A. Polyvyanyy. Process mining meets causal machine learning: Discovering causal rules from event logs. In B. F. van Dongen, M. Montali, and M. T. Wynn, editors, *Proc. of ICPM 2020*, pages 129–136. IEEE, 2020.
2. M. Camargo, M. Dumas, L. García-Bañuelos, I. Mahdy, and M. Yerokhin. Discovering business process simulation models in the presence of multitasking and availability constraints. *Data and Knowledge Engineering*, 134, July 2021.
3. M. Camargo, M. Dumas, and O. González. Automated discovery of business process simulation models from event logs. *Decis. Support Syst.*, 134:113284, 2020.
4. M. Camargo, M. Dumas, and O. G. Rojas. Discovering generative models from event logs: Data-driven simulation vs deep learning. *CoRR*, abs/2009.03567, 2020.
5. M. de Leoni, M. Dees, and L. Reulink. Design and evaluation of a process-aware recommender system based on prescriptive analytics. In *Proc. of ICPM 2020*, pages 9–16. IEEE, 2020.
6. M. Dumas, M. L. Rosa, J. Mendling, and H. A. Reijers. *Fundamentals of Business Process Management, Second Edition*. Springer, 2018.
7. N. Martin, L. Pufahl, and F. Mannhardt. Detection of batch activities from event logs. *Inf. Syst.*, 95:101642, 2021.
8. M. Pourbafrani, S. J. van Zelst, and W. M. P. van der Aalst. Supporting automatic system dynamics model generation for simulation in the context of process mining. In *Proceedings of BIS 2020*, pages 249–263. Springer, 2020.
9. M. Reichert and B. Weber. *Enabling Flexibility in Process-Aware Information Systems - Challenges, Methods, Technologies*. Springer, 2012.
10. Y. Si, V. Chan, M. Dumas, and D. Zhang. A petri nets based generic genetic algorithm framework for resource optimization in business processes. *Simul. Model. Pract. Theory*, 86:72–101, 2018.
11. S. Suriadi, M. T. Wynn, J. Xu, W. M. P. van der Aalst, and A. H. M. ter Hofstede. Discovering work prioritisation patterns from event logs. *Decis. Support Syst.*, 100:77–92, 2017.
12. F. Tao, H. Zhang, A. Liu, and A. Y. C. Nee. Digital twin in industry: State-of-the-art. *IEEE Trans. Ind. Informatics*, 15(4):2405–2415, 2019.
13. N. Tax, I. Verenich, M. L. Rosa, and M. Dumas. Predictive business process monitoring with LSTM neural networks. In *Proc. of CAiSE 2017*, pages 477–492. Springer, 2017.
14. W. M. P. van der Aalst. Business process simulation survival guide. In *Handbook on Business Process Management 1, Introduction, Methods, and Information Systems, 2nd Ed*, pages 337–370. Springer, 2015.