Note: All topics are available to both bachelor's and master's students.

Topic 1: Process Pattern Discovery in Process Mining

Process pattern discovery in process mining focuses on identifying recurring patterns or sequences of activities within business processes. Instead of solely examining individual events or activities, this approach aims to uncover common structures or behaviors that occur across multiple process instances. By recognizing these patterns, organizations can gain insights into the regularities and deviations in their processes, leading to improved understanding and potential optimizations.

This seminar thesis aims to provide an overview of process pattern discovery in process mining, emphasizing its significance and methodologies. The participant should explore recent literature on this topic, with a focus on techniques and algorithms used for pattern identification in event logs, event streams, and object-centric event data. Additionally, the thesis should address practical applications of process pattern discovery, challenges, and potential future research directions.

Literature

[1] Vazifehdoostirani, M. et al. (2023). Interactive Multi-Interest Process Pattern Discovery. International Conference on Business Process Management, 303-319.
[2] Fang, W., Zhang, Q., Sun, J., Wu, X. (2020). Mining High-Quality Patterns Using Multiobjective Evolutionary Algorithm. IEEE Transactions on Knowledge and Data Engineering, 34(8), 3883–3898.
[3] Acheli, M., Grigori, D., Weidlich, M. (2021). Discovering and Analyzing Contextual Behavioral Patterns from Event Logs. IEEE Transactions on Knowledge and Data Engineering, 34(12), 5708–5721.

Topic 2: Automated Process Improvement

The goal of business process management (BPM) is to continuously improve a running business process, hence supporting the successful operation of an organization. Given the increasing digitalization of processes and the advance of analytical techniques, such as process mining, many BPM functions have been automated in the last years. However, it is still difficult to identify concrete improvement potentials for a business process in a completely automated way, which is why this BPM task is still mainly conducted manually.

The advent of Large Language Models (LLMs), such as GPT-4, offers new opportunities for automating the improvement of business processes. LLMs offer advanced capabilities for natural language processing, which includes a certain understanding of business processes and related documents, such as textual descriptions. These capabilities, together with the analytical capabilities of process mining, might enable organizations to automatically find and implement process improvements.
The goal of this seminar thesis is to provide an overview of current approaches for automated business process improvement, with a focus on the potentials of LLMs.

Literature:

[1] Chapela-Campa, D., & Dumas, M. (2023). From process mining to augmented process execution. Software and Systems Modeling, 1-10.
[2] van Dun, C., Moder, L., Kratsch, W., & Röglinger, M. (2023). ProcessGAN: Supporting the creation of business process improvement ideas through generative machine learning. Decision Support Systems, 165, 113880.
[3] Kampik, T., et al. (2023). Large Process Models: Business Process Management in the Age of Generative AI. arXiv preprint arXiv:2309.00900.

Topic 3: Exploring System-level Business Process Traits

Business processes, as examples of complex systems, can be characterized by traits that provide insights into the properties of the process as a whole. These system-level traits, including resilience, flexibility, stability, dynamics, and complexity, offer a nuanced understanding of how business processes operate within the larger organizational context. Resilience assesses the system's ability to recover from disruptions, ensuring continuity in operations. Flexibility explores how processes can adapt to changing circumstances, promoting agility and responsiveness. Stability focuses on the consistent and reliable execution of processes under normal conditions, contributing to operational reliability. Dynamics involve studying how processes evolve over time, capturing transitions and changes in behavior. Complexity delves into the intricacy and sophistication of processes, considering factors such as task interdependencies and decision-making structures.

The seminar thesis aims to provide an overview of the analytical approaches and models used to assess and quantify the resilience, flexibility, stability, dynamics, complexity, and other system-level process characteristics of business processes. Participants are expected to explore recent state of the art on business process traits, identify relevant definitions, methodologies, and critically evaluate their applicability in different organizational contexts.

Literature

[1] Vidgof, M., Wurm, B., & Mendling, J. (2023). The Impact of Process Complexity on Process Performance: A Study using Event Log Data. International Conference on Business Process Management, 413-429.
[2] Wuyts, B. et al. (2023). DyLoPro: Profiling the Dynamics of Event Logs. International Conference on Business Process Management, 146-162.

Topic 4: Object-centric Process Modeling

Object-centric process mining shifts the focus away from the classic case notion to the objects manipulated within a business process. Rather than emphasizing events and activities in the control flow, this approach centers on objects involved in the process, such as customers, products, or documents. This provides a more intuitive representation of processes, offering a clearer understanding of the flow and interactions among objects. However, it also makes the process much more complex and therefore harder to represent in the form of a two-dimensional process model.

The goal of this seminar thesis is to provide an overview of the models used to represent object-centric processes. Therefore, the participant should identify recent papers on object-centric process mining and analyze them with regard to the applied modeling formalism. Ideally, the seminar thesis should also identify research gaps, e.g., with regard to the user-friendliness and understandability of the currently used formalisms.

Literature:

[1] van der Aalst, W., & Berti, A. (2020). Discovering object-centric Petri nets. Fundamenta informaticae, 175(1-4), 1-40.
[2] Fahland, D. (2022). Process mining over multiple behavioral dimensions with event knowledge graphs. In: Process Mining Handbook, 274-319.
[3] Berti, A., & van der Aalst, W. (2023). OC-PM: analyzing object-centric event logs and process models. International Journal on Software Tools for Technology Transfer, 25(1), 1-17.

Topic 5: Usage of Synthetic Data in Causal Process Mining

A growing number of research in the area of process mining uses causal inference (sometimes referred to as causal discovery) approaches to go beyond correlations and detect true causal relationships between features of processes. This has major advantages over conventional correlative approaches because the detection of causes can be considered actually causal.

To evaluate such causal approaches and their ability to detect true causes, researchers require data for which it is know what the actual causes for problems in the process are, often referred to as ground truth. However, since such kind of data with a causal ground truth is very rare. Therefore, researchers sometimes produce own, synthetic data in which they create causes with according ground truth. Then, they can evaluate whether approaches can re-discover this ground truth.

The goal of this seminar thesis is to collect and analyze approaches that generated synthetic data in order to evaluate causal process mining approaches. The applicant should identify 1) how the synthetic data is created, 2) what kind of data is created, 3) what kind of causes are inserted into the data, and 4) what kind of causal dependencies are detected by the corresponding techniques. The participant should finally discuss requirements for a tool that can generate synthetic data for all found applications.

Literature:
[1] Qafari, M., & van der Aalst, W. (2020). Root cause analysis in process mining using structural equation models. Business Process Management Workshops, 155-167.
[2] Leemans, S., & Tax, N. (2022). Causal reasoning over control-flow decisions in process models. International Conference on Advanced Information Systems Engineering, 183-200.

Topic 6: Inter-case Effects in Process Mining

Most existing process mining techniques assume that each case instance is a separate process execution, whose progression is only influenced by case-specific features, such as attributes or control flow decisions. However, this isolated view on cases does not account for potential inter-case effects, i.e., changes in the control flow of one case that are influenced by what happens in other cases. These effects are likely to occur because different cases of the same process are inherently connected by resources, which are typically responsible for the execution of an entire process instead of a single case.

The goal of this seminar thesis is to identify techniques that consider inter-case effects in processes. The applicant should identify 1) what kind of inter-case effects are considered, 2) how they are measured (and turned into features), and 3) which application these inter-case features have (discovery, prediction, …). Finally, the participant should discuss how inter-case features can be categorized and used in approaches.

If the number of publications in this field is too large, a specification will be done by the participant in cooperation with the supervisor before the first milestone.

Literature:
[1] Grinvald, A., Soffer, P., & Mokryn, O. (2021). Inter-case properties and process variant considerations in time prediction: A conceptual framework. International Conference on Business Process Modeling, Development and Support, 96-111.
[2] Senderovich, Arik, et al. (2017). Intra and inter-case features in predictive process monitoring: A tale of two dimensions. International Conference on Business Process Management, 306-323.

Topic 7: Declarative Process Discovery

Process discovery is a critical component of process mining, involving the analysis of event logs from information systems to construct a comprehensive model of an organization's business processes. The typical output of a process discovery algorithm is a formal process model that explicitly describes all possible execution patterns of activities in a process.

Declarative process discovery instead focuses on analyzing business processes to identify and understand the set of rules and constraints that dictate how these processes are executed. This approach is particularly useful in environments where flexibility and adaptation to changing circumstances are critical, such as the healthcare sector. Declarative process models are essentially collections of temporal or logic constraints that provide a framework for capturing such flexible processes.

The aim of this seminar thesis is to investigate the current state of the art of declarative process discovery methods. This includes exploring various algorithms and techniques used to extract declarative models from event logs. The participant should also identify the main application areas and assess the challenges faced in declarative process discovery.

Literature:

[1] Maggi, F., Mooij, A., & van der Aalst, W. (2011). User-guided discovery of declarative process models. IEEE Symposium on Computational Intelligence and Data Mining, 192-199.
[2] Pesic, M., Schonenberg, H., & van der Aalst, W. (2007). DECLARE: Full support for loosely-structured processes. IEEE International Enterprise Distributed Object Computing Workshop. 287.
[3] Rovani, M., Maggi, F., de Leoni, M., & van der Aalst, W. (2015). Declarative process mining in healthcare. Expert Systems with Applications, 42(23), 9236-9251.

Topic 8: Business Process Simulation Models

Business process simulation involves creating and evaluating models that mimic the dynamics of real-world business processes. Unlike traditional static models, simulation models allow for the dynamic representation of processes, considering factors such as resource utilization, variability, and performance over time. These models serve as valuable tools for decision-makers to test different scenarios, assess the impact of changes, and optimize processes before implementation.

The objective of this seminar thesis is to provide a comprehensive overview of the development and evaluation of business process simulation models. Participants are expected to explore recent literature on methodologies, techniques, and tools used in creating these models. Additionally, the seminar thesis should delve into the evaluation criteria for assessing the effectiveness and accuracy of simulation models in capturing the intricacies of business processes.

Literature

[1] López-Pintado, O, & Dumas, M. (2023). Discovery and Simulation of Business Processes with Probabilistic Resource Availability Calendars. International Conference on Process Mining, 1-8.
[2] Knopp, B., Pourbafrani, M. & van der Aalst. W. (2023) Discovering Object-Centric Process Simulation Models. International Conference on Process Mining, 81-88.
[3] Chapela-Campa, D. et al. (2023). Can I Trust My Simulation Model? Measuring the Quality of Business Process Simulation Models. arXiv preprint arXiv:2303.17463.

Topic 9: Stochastic Process Mining

Stochastic process mining involves analyzing and modeling business processes with inherent variability and uncertainty. Unlike traditional process mining, which assumes deterministic behavior, stochastic process mining acknowledges the probabilistic nature of events in real-world processes. This approach incorporates randomness, variability, and timing uncertainties, providing a more realistic representation of complex systems. Key concepts in stochastic process mining include the utilization of stochastic Petri nets, Markov models, and other probabilistic formalisms to capture and analyze the dynamics of processes. By considering the stochastic nature of events, practitioners gain insights into the inherent variability of process executions, enabling better decision-making and performance optimization.

The goal of this seminar thesis is to provide an overview of stochastic process mining and how it differs from traditional process mining. Therefore, the participant should identify recent papers on stochastic process mining, with a focus on the discovery of different process perspectives.

Literature:
[1] Gal, A. (2023). Everything there is to Know about Stochastically Known Logs. International Conference on Process Mining, xvii-xxiii.
[2] Burke, A., Leemans, S., & Wynn, M. (2020). Stochastic process discovery by weight estimation. International Conference on Process Mining, 260-272.
[3] Mannhardt, F., Leemans, S., Schwanen, C., & de Leoni, M. (2023). Modelling Data-Aware Stochastic Processes-Discovery and Conformance Checking. International Conference on Applications and Theory of Petri Nets and Concurrency, 77-98.

Topic 10: Properties of Event Logs

Process mining involves analyzing event logs that capture the behavior of business processes within enterprise information systems in order to gain insights into and eventually improve those processes. Event logs hence serve as data input for almost every process mining algorithm. However, not all event logs are the same. Depending on the underlying process and multiple other factors, they can differ substantially from one another, for example with regard to their completeness, amount of noise, size, or variability. These properties of events logs may in turn impact the performance of some process mining algorithms, for example for discovery or process prediction. This means that a proper characterization of event logs is important to select the most appropriate process mining algorithms.

The goal of this seminar thesis is to curate a selection of properties that can be used to characterize and compare event logs in process mining. This should include both a general intuition of the property as well as an operationalization in terms of a proper measurement.

Literature:
[1] vanden Broucke, S., Delvaux, C., Freitas, J., Rogova, T., Vanthienen, J., & Baesens, B. (2014). Uncovering the relationship between event log characteristics and process discovery techniques. Business Process Management Workshops, 41-53.
[2] Janssenswillen, G., Jouck, T., Creemers, M., & Depaire, B. (2016). Measuring the quality of models with respect to the underlying system: An empirical study. International Conference on Business Process Management, 73-89.
[3] Kratsch, W., Manderscheid, J., Röglinger, M., & Seyfried, J. (2021). Machine learning in business process monitoring: a comparison of deep learning and classical approaches used for outcome prediction. Business & Information Systems Engineering, 63, 261-276.

Topic 11: Explainable Process Mining

Extracting insights from event logs in information systems through process mining often involves complex algorithms and models. While these models are powerful in analyzing and optimizing business processes, their inherent complexity can make it challenging for users to understand how specific insights or recommendations are derived. This lack of transparency can be a significant barrier, especially in contexts where trust and clarity in decision-making are paramount.

The goal of this seminar thesis is to examine existing research that focuses on making process mining algorithms more transparent and understandable to users. This encompasses both retrospective explainability methods that can be implemented in conjunction with existing process mining techniques, and inherently explainable process mining approaches that were designed with understandability in mind. The participant should evaluate how the identified explainability approaches work (e.g., feature attributions in machine learning techniques) and how they are being integrated into different process mining techniques, such as discovery and conformance checking.

Literature:
[1] Stierle, M., Brunk, J., Weinzierl, S., Zilker, S., Matzner, M., & Becker, J. (2021). Bringing Light Into the Darkness - A Systematic Literature Review on Explainable Predictive Business Process Monitoring Techniques. European Conference on Information Systems, 1175.
[2] Adams, J., van Zelst, S., Quack, L., Hausmann, K., van der Aalst, W., & Rose, T. (2021). A Framework for Explainable Concept Drift Detection in Process Mining. International Conference on Business Process Management, 400-416.

Topic 12: Categorization of Process Deviations

Ensuring the compliance of business processes is of utmost importance for organizations. To detect compliance violations, managers can apply conformance checking. These techniques compare recorded process executions in an event log with their intended state in form of a process model. They return a set of deviations that occurred in the event log.

However, not all deviations can be considered equally severe. In some cases, a deviation might be considered an exception that is still compliant (e.g., although typically required, a hospital patient might not be registered at the front desk before heading to an operating room if (s)he is in a life-threatening condition). Therefore, it is of importance to analyze what kind of deviation is detected, including how severe it is, whether it was an exception, whether it affects the process outcome, or whether it indicates inefficiencies.

The goal of this seminar thesis is to identify frameworks and techniques that categorize deviations. The applicant should identify 1) what kind of deviation categories are typically used, 2) whether the categories are characterized by any attributes (e.g., severity), and 3) whether and if so, how they are detected in reality.

Note that the relevant literature is not limited to process mining specific literature but also includes Business Process Management Literature.

Literature:
[1] Swinnen, J., et al. (2011). A process deviation analysis – a case study. Business Process Management Workshops: Business Process Management Workshops, 87-98.
[2] Laghmouch, M., Jans, M., & Depaire, B. (2020). Classifying process deviations with weak supervision. International Conference on Process Mining, 89-96.