Topic 1: Conformance Checking with Attributes

Conformance checking is one of the main tasks of process mining. It comprises techniques for checking the relation between a designed process model and the real-life behavior of a process, with the goal of identifying and analyzing deviations between them. State-of-the-art conformance checking approaches have become computationally very efficient and robust to many kinds of process behavior. However, most of those approaches focus only on the control-flow perspective, i.e., only take the activities from the event log into account. This might limit their applicability in many practical applications, where the order of the activities might not be the (sole) relevant factor for the conformance of a process. Instead, other process attributes like cost or time might play a major role.

The goal of this seminar thesis is to identify approaches to conformance checking that deliberately include attributes for determining the conformance between a log and a model. Therefore, the participant should conduct a structured literature review. In a second step, the identified approaches should be analyzed with regard to (1) the attributes they consider, (2) the assumptions they make, and (3) the computational techniques that they use. The participant should also analyze the technical, computational, and conceptual challenges that conformance checking with attributes holds and the research gaps that still exist in this area.

Topic 2: Conformance Checking in Practice

Conformance checking is one of the main tasks of process mining. It comprises techniques for checking the relation between a designed process model and the real-life behavior of a process, with the goal of identifying and analyzing deviations between them. State-of-the-art conformance checking approaches have become computationally very efficient and robust to many kinds of process behavior. However, those approaches are typically designed by researchers in a generic way, meaning that they are not geared towards a specific use case. The business value of conformance checking remains to be identified by the organizations that want to use them.

The goal of this seminar thesis is to compile a catalogue of practical use cases for conformance checking. Therefore, the participants should conduct a structured literature and identify case studies and practice reports, where a certain conformance checking technique was used for a specific purpose. In compiling the catalogue, the participant should distinguish between domain-specific and generic use cases, clarify the assumptions about available data, and specify the concrete business value that the use case provides.

Topic 3: Process Mining meets Visual Analytics

Process mining is a research field that focuses on analyzing business processes that are executed on IT systems, based on their behavior as recorded in an event log. With the growing availability of event log data and the advanced maturity of tools and algorithms, process mining has recently also gained widespread industry adoption, with many organizations implementing tools or conducting projects as a central part of their business process management strategy. This means that process mining results are presented to business analysts and managers, who have little expertise in process analytics, but have to make decisions based on their results. Visualizing these results in a user-centric way is therefore a vital part of establishing process mining in organizations.

Visual analytics is a research field that focuses on facilitating analytical reasoning by means of data analytics techniques and interactive visual interfaces. It combines multiple disciplines including computer science, cognitive psychology, and graphic design in order to visualize complex situations and assists humans in understanding and advancing them.

The goal of this master thesis is to identify application potentials for visual analytics in process mining. Therefore, the participant should first define both concepts and analyze process mining from a visual analytics standpoint. In a second step, the participant should conduct a structured literature review in order to identify works that already combine the two fields. Because the literature in this area is suspected to be sparse, the participant should as a third step analyze state-of-the-art process mining tools (such as Celonis, Disco, ProM, etc.) to analyze how they could benefit from visual analytics.

Topic 4: Designing Dashboards for Business Process Management

A dashboard is a graphical user interface which provides its users with a high-level overview of all relevant key performance indicators (KPIs) of a particular objective or business process. It is considered as a form of data visualization. In today's increasingly data-driven business environments, dashboards play a central role for managers' ability to monitor and control their organizations. Therefore, it is vitally important that they are designed in a user-centric and effective manner.

The goal of this seminar thesis is to derive a set of principles for designing a dashboard for process-centric data. Therefore, the participant should conduct a structured literature review in order to identify dashboards that were designed in a business process management context and analyze commonalities between them in order to derive recommendation for the design of future dashboards. In a second step, the participant should conduct a second structured literature review to identify general principles for dashboard designs that can be applied to process-centric data. The final thesis should combine results from both studies and serve as the basis for designing a new process-centric dashboard.

Topic 5: Natural Language Processing Methods in Process Mining

Text and process data share some common characteristics: both can be interpreted as sequences of non-numeric tokens, both are constrained by external "rules" in the form of grammar and process logic, and the frequency distribution of events in event logs tends to be very similar to the distribution of words in documents. Motivated by these commonalities, many researchers have adopted methods from the field of Natural Language Processing (NLP) to solve problems in process mining. Examples of successful applications include next-activity-prediction with Recurrent Neural Networks and process representation learning with Word2vec.

The objective of this seminar thesis is to conduct a structured literature review to identify research papers that apply NLP methods to process traces recorded in event logs. In particular, the review should detail how each method works, which specific Process Mining challenge(s) it addresses, and how it has been adapted to be applicable to process data.

Topic 6: Detecting Fraud with Process Mining

Fraudulent behavior is a considerable problem in modern businesses and has significant negative consequences for customers, employees, investors, and creditors. Corporate fraud schemes can be highly intricate, and traditional fraud detection as performed by forensic specialists in auditing firms and government agencies is not always successful in discovering them before the damage is done – prominent examples are the recent Wirecard financial scandal or Volkswagen's use of defeat devices to cheat emission standards.

As companies digitalize their process landscape, an ever-growing number of stakeholder activities can be traced in the event logs created in day-to-day operations. process mining techniques can be used to analyze these logs and find fraudulent cases in historical data, and even to derive models that can automatically flag suspicious deviations in future process executions. process mining thus has the potential to support the audit process for both minor and major cases of fraud.

The objective of this seminar thesis is to identify existing approaches to process-mining-based fraud detection by conducting a structured literature review. Besides detection approaches for internal and external fraud (e.g., by customers in online shops), the review can also include approaches for other intentional process deviations that do not meet the legal definition of fraud (like employee work-arounds).

Topic 7: Metrics for Software Usability

The usability of an Information System (IS) is a key factor in designing and selecting business application software. If an IS has a high usability, it is easy to use, easy to learn, leads to a satisfactory outcome for the user, is efficient to use, and easy to memorize. Because usability has been shown to be a clear competitive advantage for any software, the field of usability engineering has dedicated itself to researching how a high degree of software usability can be achieved. In addition to qualitative methods for assessing usability (such as think-aloud studies), quantitative methods for measuring software usability have received considerable interest. Those methods are typically based on so-called software usability metrics, which measure the usability of a certain software.

The goal of this seminar thesis is to compile a list of software usability metrics, which are used in state-of-the-art IS usability engineering. Therefore, the participant should conduct a structured literature review, with a focus on business software applications. For each usability metric, the objective, formula (if applicable), required data, and any other type of relevant property should be stated. If necessary, the resulting list should distinguish between mobile devices and PCs, focusing on the latter.

Topic 8: Task Discovery Techniques for Robotic Process Automation

Robotic process automation (RPA) aims to automate repetitive business routines by training bots to emulate human behavior. The first step in an RPA project is to record all steps that human actors perform when executing a routine. This is usually implemented by enhancing an information system with logging software that records low-level user interface interactions such as clicking a button or writing a string into a text field. The resulting user interaction (UI) logs often span multiple days or weeks and do not include task labels, i.e., there is no indicator for which of the recorded user actions belong to the same task. Traditionally, these labels are provided by process experts, which is time-consuming and error-prone. A promising alternative to manual annotation is automated task discovery, in which data mining methods like clustering and supervised classification are applied to segment the activities recorded in UI logs into tasks.

The goal of this seminar thesis is to investigate the effectiveness and applicability of these automated approaches. Therefore, the participant should conduct a structured literature review to find techniques that can discover distinct tasks in UI logs. The techniques found should be analyzed and compared with regards to their capabilities and requirements (e.g., training data or log pre-processing). While the focus of the review is on approaches that enable RPA, more general techniques for UI log segmentation can also be considered.