

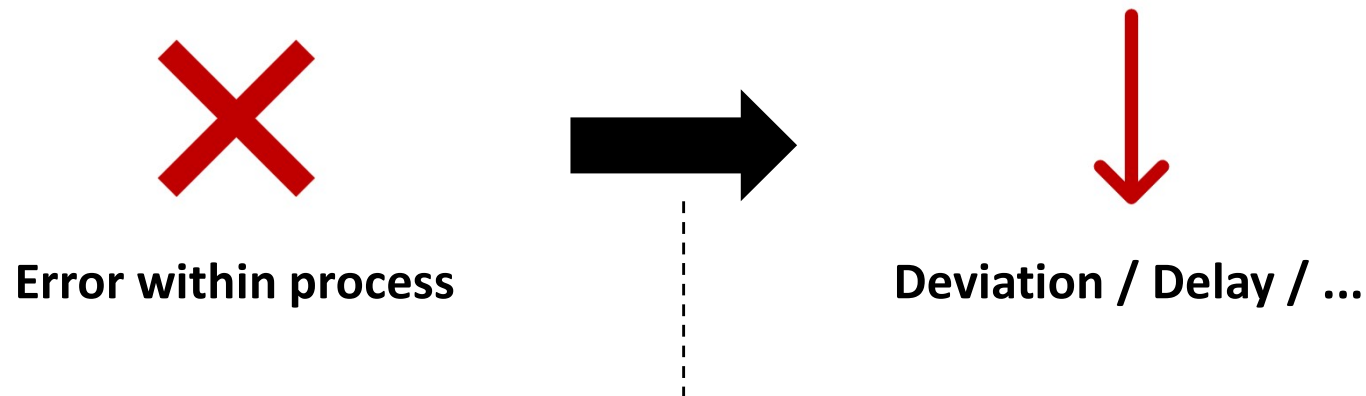
Simulating Causes in Event Logs with Ground Truth

Michael Grohs & Jana-Rebecca Rehse



Causal Inference in Process Mining

Goal:



- **Establishment of causes** is typically done by **correlating process characteristics**
- However, causes should be in fact causal and **not mere correlations**
- Therefore, more and more approaches use **causal inference in process mining**

Problem:



True causes in real-life event logs **are not known**
-> researchers rely on simulations of causes with ground truth for
evaluation purposes

Project Goals

Build a tool that simulates certain errors (concretely, conformance violations) in the process

- Errors should be **measurable within the event log** (i.e., deviations from a process model)
- Errors should be **attributable to variations in other attributes**
- Tool should incorporate **choice of noise level**
- General idea: **“If the resource ABC is involved into a process instance, the likelihood that error X occurs is increased by 50%”**

Project Requirements

- Duration of the project: 6 Months
- Minimum/maximum number of participants: 2-4
- Prerequisites: **(IS515 or IS514 or IE692) and (Programming skills in Python and some GUI dev skills (e.g., JS Frameworks))**
- Interest in Process Mining
- Applicable for Business Informatics and MMDS

Questions: michael.grohs@uni-mannheim.de

Illustration of Inserted Causes

Activity „Delete Order“ is skipped in various process instances

The **reason for that is a higher delivery time** which causes the process instance to deviate

A high delivery time increases the **likelihood of the deviation by 30%**

