LTAT.05.015

Business Process Mining

Lecture 5: Conformance Checking

Marlon Dumas

Professor of Information Systems @ University of Tartu
Co-founder @ Apromore
# Course Outline

<table>
<thead>
<tr>
<th>Section</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Introduction</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Process performance measurement &amp; dashboards</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Process mining</strong></td>
<td></td>
</tr>
<tr>
<td>• Automated process discovery</td>
<td></td>
</tr>
<tr>
<td>• <strong>Conformance checking</strong></td>
<td></td>
</tr>
<tr>
<td>• Performance mining</td>
<td></td>
</tr>
<tr>
<td>• <strong>Variant analysis</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Simulation and “what-if” process mining</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Data extraction &amp; preparation for process mining</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Predictive process monitoring</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Business case analysis &amp; project management for process mining</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Trends: prescriptive monitoring, causal process mining, robotic process mining</strong></td>
<td></td>
</tr>
</tbody>
</table>
Process Mining

Automated Process Discovery

Performance Mining

Conformance Checking

Variants Analysis

Business rules / normative model

Enhanced process model

discovered process model

event log

Difference diagnostics

event log'

✅ / ✗
### Typical Process Mining Analysis Workflow

**Start with automated process discovery (exploratory phase)**

- Perform sanity checks to determine if we have the right dataset (e.g. are there are incomplete cases?)
- Understand the structure of the process, identify locations of rework/repetition, key decision points and parallel branches

**Depending on the emphasis of the project, use either:**

- Compliance checking to find undesired deviations and exceptions
- Performance mining to identify bottlenecks and wastes in the process

**Use variant analysis to dig deeper. Depending on the emphasis of the project:**

- Use variant analysis to analyze non-compliant versus compliant cases
- And/or to understand performance differences between two countries, regions, types of customers, etc
Given an event log and a set of **business rules** or a **process model**, find, describe, and measure the impact of **differences** between the rules/model and the log.

<table>
<thead>
<tr>
<th>No. of Instances</th>
<th>Log Traces</th>
</tr>
</thead>
<tbody>
<tr>
<td>1207</td>
<td>ABDEA</td>
</tr>
<tr>
<td>145</td>
<td>ACDGHFA</td>
</tr>
<tr>
<td>56</td>
<td>ACGDHFA</td>
</tr>
<tr>
<td>23</td>
<td>ACHDFRA</td>
</tr>
<tr>
<td>28</td>
<td>ACDHFA</td>
</tr>
</tbody>
</table>

Log: ABDEA, ACDGHFA, ACGDHFA, ACHDFRA, ACDHFA

Rules, process model: ≠

**Conformance checking**
Types of Conformance Checking

1. Rule-Based Conformance Checking
   • Checking that a process follows certain rules

2. Model-Based Conformance Checking
   • Comparing the behaviour captured in the process model against the one in the event log

3. Exceptional Behavior Analysis
   • Spotting anomalous behavior
Rule-Based Conformance Checking

Checking whether regulations and control measures embedded in processes are followed in practice.

Common types of compliance rules:

**Flow constraints**

Certain activities may be mandatory e.g. a specific approval to proceed with an application.

**SLA constraints**

Customer agreements on service quality define timeframes at the case, milestone or activity transition level.

**Resource constraints**

Certain activities must be executed by different employees, e.g. creation and approval of an invoice.
Conformance Checking Levels

Accordingly, we can identify the following *compliance checking levels*:

- **Flow**: have all mandatory activities been performed? Have activities been performed in the right order and with the right frequency?
- **Temporal**: has a given timeframe between activities, process milestones or for the end-to-end process been adhered to?
- **Resource**: did everyone obey the segregation of duties?
- **Exception**: explore highly-infrequent behavior to identify possible deviations from the standard process

**Procedure:**

1. Define compliance rules (based on attribute-value pairs, activity orders or performance targets)
2. Apply one or more filters to keep cases that violate the rules in question
3. If needed, switch perspective to highlight non-compliance issues (e.g. show resources that violate a segregation of duties principle)
4. Assess impact on process performance
## Conformance Checking: Analysis Template

<table>
<thead>
<tr>
<th>What?</th>
<th>How?</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Flow compliance checking</strong></td>
<td>Identify skipped activities in the process map or BPMN model</td>
</tr>
<tr>
<td>1. Check skipped mandatory activities</td>
<td>Visualize process map with the min/max frequency metrics</td>
</tr>
<tr>
<td>2. Check forbidden repetitions.</td>
<td>Visualize min/max statistics in Activity tab of the dashboard (EE)</td>
</tr>
<tr>
<td>3. Check activity exclusion or co-occurrence relations</td>
<td>Use Path filter (directly-follows or eventually-follows)</td>
</tr>
<tr>
<td><strong>Temporal compliance checking</strong></td>
<td>Use performance filter (to check for end-to-end SLA violations)</td>
</tr>
<tr>
<td>Check temporal constraints violations</td>
<td>Use Path filter (eventually-follows) with a duration constraint to identify violations of temporal constraints between two activities</td>
</tr>
<tr>
<td><strong>Resource compliance checking</strong></td>
<td>Use Path filter (eventually-follows) and require the same Resource (or different resources)</td>
</tr>
<tr>
<td>Check four-eyes principle or same-case-hander principle</td>
<td>Use the Attribute filter and search with “Activity” as the primary attribute and “Resource” as secondary attribute to check if a resource performed an activity they were not supposed to do.</td>
</tr>
<tr>
<td>Check if an activity has been performed by a resource who should not do so.</td>
<td></td>
</tr>
<tr>
<td><strong>Model-to-log conformance checking</strong></td>
<td>Use a “model-to-log” conformance checker</td>
</tr>
<tr>
<td>Identify deviations between the event log and the process model and analyze their frequency and impact.</td>
<td>Or discover a model from a log and compare it manually with the manually designed process model</td>
</tr>
<tr>
<td><strong>Exception analysis</strong></td>
<td>Use abstraction slider and invert the order in which arcs/nodes are removed (remove most frequent arcs/nodes)</td>
</tr>
</tbody>
</table>
Flow Compliance Checking

Mandatory tasks: Have all mandatory tasks been completed?

To check whether all your tasks have been executed at least once, change the frequency view to “Min.” in the visualization settings of your Process Discoverer.

The numbers in your activities now indicate the minimum amount of times it has been completed in all process instances.
Temporal Compliance Checking

Prescribed time frame: Has the given time frame been adhered to?

Select the “Path” filter and define the steps that must be performed in a defined time frame. Use “Time interval” to then set your temporal figures.

Example:

Compliance rule: Waiting time between “Inform User” and “Repair (Simple)” should not exceed 1 hours.

“Time Interval” Filter: > 1 hour, so that only those cases are retained that do not comply with the compliance rule.
Performance filter can help to check for end-to-end SLA violations.

For example, retain only those cases that get completed in at most 1 hour.
Resource Compliance Checking

1 Segregation of duties: Did everyone obey the segregation of duties?

To check whether the segregation of duties has been breached, select a “Path” filter. Now, set another filter which indicates that the same resource must have completed the tasks (e.g. creating & approving an invoice)
Resource Compliance Checking

1. Unexpected resource assignments

User the Attribute filter with two attributes (primary and secondary) to retrieve cases where a given activity was performed by a given resource (or any other combination of attributes)
Model-Based Conformance Checking

Unfitting behaviour:
- Task C is *optional* (i.e. may be skipped) in the log

Additional behavior:
- The *cycle* including IGDF is not observed in the log

Event log:
- ABCDEH
- ACBDEH
- ABCDFH
- ACBDFH
- ABDEH
- ABDFH
Exception Analysis

Analyze highly infrequent behavior

Invert the “abstraction” order to see outliers. By reversing the ordering, it is no longer the most frequent cases that are displayed first, but the rarest cases. This shows how much the less frequent process deviated from the standard process.
Exception Analysis

1. Analyze highly infrequent behavior – extract most infrequent case variants

Select the most infrequent case variants in the case variants filter. You will note here a lot of the exceptions: cases aborted prematurely, rework loops, etc.
Process Mining

- Automated Process Discovery
- Performance Mining
- Conformance Checking
- Variants Analysis
- Enhanced process model
- event log
- event log'
- Business rules / normative model
- Discovered process model
- Difference diagnostics

- Tropical
- Moderate to subtropical
- Subtropical
- Oranges
- Apples
- Fruit
- Grow
- Trees
## Variant Analysis Template

<table>
<thead>
<tr>
<th>What?</th>
<th>How?</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Flow comparison</strong>&lt;br&gt;Identify notable diffs in activity flow</td>
<td>• Discover a BPMN process model from the log of each variant. Visualize them side-by-side. Look for the gateways around each task to detect for example situations where two tasks are in parallel in one variant, but sequential in the other.</td>
</tr>
</tbody>
</table>
| **Frequency & rework comparison**<br>Compare activity or handoff frequency and rework | • Compare the maps of the process variants side-by-side (as above), but using the frequency overlay instead of the duration overlay. This allows us to detect the most frequent transitions between tasks. Look also for “thick” loop-backs.  
• Use a multi-low dashboard to compare frequencies of activities, resources or other attributes. |
| **Bottleneck comparison**<br>Compare the location and magnitude of bottlenecks across variants | • Use filtering interface to retain all traces that fulfill the condition to belong to a variant (then repeat for the other).  
• Open the maps of the process variants side-by-side and compare the dependencies (arcs) with the highest waiting times. If the maps are too complex, use the abstraction slider with the abstraction metric “average duration” and the ordering from slow (right) to fast (left). Use the arc slider to retain the slowest dependencies.  
• Compare activity durations using side-by-side comparison of process maps or a multi-log dashboard (open the logs of 2+ variants in a single dashboard).  
• Consider comparing the two variants using the “resource” perspective to identify hand-offs between resources.  
• Use a multi-log animation to observe the build up of bottlenecks. |
Variant Analysis

Find differences in performance and control-flow between multiple variants of a process

A process variant is a subset of the executions of a process corresponding to a given product, customer, etc.
To start a detailed variant analysis, the event log has to be divided by a criterion.

Example: divide by case duration:
Variant 1: Short case duration (Fast cases)
Variant 2: Long case duration (Slow cases)

After the desired filter has been applied, you may each variant as a filtered log.
Variant Analysis

After you have saved your filtered logs, the actual analysis can start. There are three options to carry out a variant analysis in Apromore.

1. Visual comparison of the variants

Open both filtered logs in the Process Discoverer. After that, place them side by side in your preferred browser.

This enables a first visual comparison. Our example shows that the underlying process of the slow and fast process instances clearly differs.
Variant Analysis

After you have saved your filtered logs, the actual analysis can start. There are three options to carry out a variant analysis in Apromore: comparison in dashboard, via animation, or side-by-side process map comparison.

Comparison in the dashboard

Select the filtered logs and simultaneously open them in the performance dashboard.

All dashboard functionalities are also available when comparing multiple variants.
**Variant Analysis**

You can animate two or more logs together to compare their dynamics over time.

3 **Comparison in the animation**

Select your filtered logs as well as a BPMN model and start the animation. The animation enables us to visually identify the different movements of cases through the model.
In order to get detailed information on the temporal performance of processes, a filter has to be set in Apromore.

Select the “Performance” filter in Apromore and take a detailed look at those cases that differ significantly from the average duration.

You can find the average cycle time under “Case duration” in the Process Discoverer and in the dashboard.

Save the above- and below-average cases in a separate event log or separate process models each to compare them and to identify differences using the Variant Analysis template.
Next Week

Automated Process Discovery

Performance Mining

Conformance Checking

Variants Analysis

Enhanced process model

event log

Business rules / normative model

Discovered process model

event log'

Difference diagnostics

Valid / Invalid