Operational Support

prof.dr.ir. Wil van der Aalst

www.processmining.org
Process mining spectrum

“world”
- business processes
- people
- machines
- components
- organizations

software system
- records events, e.g., messages, transactions, etc.

(models) model
- analyzes
- specifies configures implements analyzes

event logs
- discovery
- conformance
- enhancement

supports/controls
Refined process mining framework
Business process provenance

- people
- machines
- business processes
- “world”
- organizations
- documents
- historic data
- provenance
- event logs
- current data
- “pre mortem”
- historic data
- “post mortem”
Two types of event data: post and pre mortem

• “Post mortem” event data refer to information about cases that have completed, i.e., these data can be used for process improvement and auditing, but not for influencing the cases they refer to.

• “Pre mortem” event data refer to cases that have not yet completed. If a case is still running, i.e., the case is still “alive” (pre mortem), then it may be possible that information in the event log about this case (i.e., current data) can be exploited to ensure the correct or efficient handling of this case.
Two types of models: “de jure models” and “de facto models”

- A de jure model is normative, i.e., it specifies how things should be done or handled. For example, a process model used to configure a BPM system is normative and forces people to work in a particular way.
- A de facto model is descriptive and its goal is not to steer or control reality. Instead, de facto models aim to capture reality.
• **Discover.** This activity is concerned with the extraction of (process) models.

• **Enhance.** When existing process models (either discovered or hand-made) can be related to events logs, it is possible to enhance these models.

• **Diagnose.** This activity does not directly use event logs and focuses on classical model-based analysis.
Auditing

- **Detect.** Compares de jure models with current “pre mortem” data. The moment a predefined rule is violated, an alert is generated (online).

- **Check.** The goal of this activity is to pinpoint deviations and quantify the level of compliance (offline).

- **Compare.** De facto models can be compared with de jure models to see in what way reality deviates from what was planned or expected.

- **Promote.** Promote parts of the de facto model to a new de jure model.
• **Explore.** The combination of event data and models can be used to explore business processes at run-time.

• **Predict.** By combining information about running cases with models, it is possible to make predictions about the future, e.g., the remaining flow time and the probability of success.

• **Recommend.** The information used for predicting the future can also be used to recommend suitable actions (e.g. to minimize costs or time).
Operational support: online process mining using “pre mortem” event data

known past

future

current state

detect: b does not fit the model (not allowed, too late, etc.)

predict: some prediction is made about the future (e.g., completion date or outcome)

recommend: based on past experiences c is recommended (e.g., to minimize costs)

T=10
Let us focus one time

<table>
<thead>
<tr>
<th>case id</th>
<th>trace</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>(a_{\text{start}}^{12}, a_{\text{complete}}^{19}, b_{\text{start}}^{25}, d_{\text{start}}^{26}, b_{\text{complete}}^{32}, d_{\text{complete}}^{33}, e_{\text{start}}^{35}, e_{\text{complete}}^{40}, h_{\text{start}}^{50}, h_{\text{complete}}^{54})</td>
</tr>
<tr>
<td>2</td>
<td>(a_{\text{start}}^{17}, a_{\text{complete}}^{23}, d_{\text{start}}^{28}, c_{\text{start}}^{30}, d_{\text{complete}}^{32}, c_{\text{complete}}^{38}, e_{\text{start}}^{50}, e_{\text{complete}}^{59}, g_{\text{start}}^{70}, g_{\text{complete}}^{73})</td>
</tr>
<tr>
<td>3</td>
<td>(a_{\text{start}}^{25}, a_{\text{complete}}^{30}, c_{\text{start}}^{32}, c_{\text{complete}}^{35}, d_{\text{start}}^{35}, d_{\text{complete}}^{40}, e_{\text{start}}^{45}, e_{\text{complete}}^{50}, f_{\text{start}}^{50}, f_{\text{complete}}^{55}, b_{\text{start}}^{60}, d_{\text{start}}^{62}, b_{\text{complete}}^{65}, d_{\text{complete}}^{67}, e_{\text{start}}^{80}, e_{\text{complete}}^{87}, g_{\text{start}}^{90}, g_{\text{complete}}^{98})</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>
Transition system (with start/complete)

Transition system diagram with states and transitions labeled.
Operational support: Detect
Declare specifications for detecting violations

- **Satisfied**: the LTL formula evaluates to true for the current partial trace.
- **Temporarily violated**: the LTL formula evaluates to false, however, there is a longer trace that evaluates to true.
- **Permanently violated**: the LTL formula evaluates to false for current trace and all its extensions.
Conflicting constraints

• A Declare specification is 
satisfied for a case if all of its 
constraints are satisfied.
• A Declare specification is 
temporarily violated by a case if 
for the current partial trace at 
least one of the constraints is 
violated, however, there is a 
possible future in which all 
constraints are satisfied.
• A Declare specification is 
permanently violated by a case if 
no such future exists.

Note that c1, c2, and c3 imply that e cannot 
be executed without permanently violating 
the specification.
Operational support: Predict

enterprise information system

operational support system

predicted completion date: 25-4-2011

partial trace

predictive model

event log

learn
Examples of predictions

• the predicted remaining flow time is 14 days;
• the predicted probability of meeting the legal deadline is 0.72;
• the predicted total cost of this case is 4500 euro;
• the predicted probability that activity a will occur is 0.34;
• the predicted probability that person r will work on this case is 0.57;
• the predicted probability that a case will be rejected is 0.67; and
• the predicted total service time is 98 minutes.
Annotated transition system
Collect results per state

elapsed times: [21, 21, 15, 42, ...]
remaining times: [21, 35, 58, 31, ...]
sojourn times: [2, 12, 5, 13, ...]

average remaining flow time is 42.56
Example: Predicting the Remaining Processing Time in a Municipality
Same event log but a coarser abstraction
Operational support: Recommend

- Enterprise information system
- Operational support system
- Model
- Event log

Partial trace

Recommendation:

x (85% certainty)
y (12% certainty)
z (3% certainty)

Suggestion: do x
Recommend

• **Possible recommendations:**
  - next activity;
  - suitable resource; or
  - routing decision.

• **A recommendation is always given with respect to a specific goal.**

• **Examples of goals are:**
  - minimize the remaining flow time;
  - minimize the total costs;
  - maximize the fraction of cases handled within 4 weeks;
  - maximize the fraction of cases that is accepted; and
  - minimize resource usage.
Relation between prediction and recommendation

possible next state

current state

prediction

a_1

a_2

a_k

...
Process mining spectrum

- information system(s)
  - people
  - machines
  - business processes
  - organizations

“world”

- current data
- historic data
- documents

- historic data
- rules
- control-flow

- de jure models
  - control-flow
  - data/rules
  - resources/organization

- de facto models
  - control-flow
  - data/rules
  - resources/organization

- event logs
  - "pre mortem"
  - "post mortem"

- provenance
- event logs
- navigation
- auditing
- cartography

- explore
- predict
- recommend
- detect
- check
- compare
- promote
- discover
- enhance
- diagnose

- models

- "pre mortem"
- "post mortem"
Three Key Observations
#1 Alignments are essential!

- conformance checking to diagnose deviations
- squeezing reality into the model to do model-based analysis

<table>
<thead>
<tr>
<th>a</th>
<th>c</th>
<th>b</th>
<th>d</th>
<th>t5</th>
<th>t7</th>
<th>f</th>
<th>h</th>
<th>t10</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>c</td>
<td>b</td>
<td>d</td>
<td>τ</td>
<td></td>
<td>f</td>
<td>h</td>
<td>t10</td>
</tr>
</tbody>
</table>

move on model

move on model (harmless)

move on log

move on model
#2 Models are like the glasses required to see and understand event data!
More and more information about business processes is recorded by information systems in the form of so-called "event logs". Despite the omnipresence of such data, most organizations diagnose problems based on fiction rather than facts. Process mining is an emerging discipline based on process model-driven approaches and data mining. It not only allows organizations to fully benefit from the information stored in their systems, but it can also be used to check the conformance of processes, detect bottlenecks, and predict execution problems.

Wil van der Aalst delivers the first book on process mining. It aims to be self-contained while covering the entire process mining spectrum from process discovery to operational support. In Part I, the author provides the basics of business process modeling and data mining necessary to understand the remainder of the book. Part II focuses on process discovery as the most important process mining task. Part III moves beyond discovering the control flow of processes and highlights conformance checking, and organizational and time perspectives. Part IV guides the reader in successfully applying process mining in practice, including an introduction to the widely used open-source tool ProM. Finally, Part V takes a step back, reflecting on the material presented and the key open challenges.

Overall, this book provides a comprehensive overview of the state of the art in process mining. It is intended for business process analysts, business consultants, process managers, graduate students, and BPM researchers.

Features and Benefits:
- First book on process mining, bridging the gap between business process modeling and business intelligence.
- Written by one of the most influential and most-cited computer scientists and the best-known BPM researcher.
- Self-contained and comprehensive overview for a broad audience in academia and industry.
- The reader can put process mining into practice immediately due to the applicability of the techniques and the availability of the open-source process mining software ProM.