Combining Workflow Management and Process Mining

CSIRO/Hobart, 2-10-2007

Prof.dr.ir. Wil van der Aalst
Eindhoven University of Technology,
P.O. Box 513, 5600 MB Eindhoven, The Netherlands
w.m.p.v.d.aalst@tue.nl
Outline

1. Overview Process Aware Information Systems
2. Workflow Patterns (short)
3. Process Verification (short)
4. Process Mining (long)
5. Conclusion

The work of many people!

Thanks to Ton Weijters, Boudewijn van Dongen, Ana Karla Alves de Medeiros, Anne Rozinat, Christian Günter, Eric Verbeek, Ronny Mans, Minseok Song, Laura Maruster, Huub de Beer, Peter van den Brand, Jan Mendling, Andriy Nikolov, Jianmin Wang, Lijie Wen, Irene Vanderfeesten, Mariska Netjes, Steffi Rinderle, Walid Gaaloul, Gianluigi Greco, Antonella Guzzo, etc. etc.
Overview Process Aware Information Systems (PAIS)
Software systems are the mirror image of the “world”

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

supports/controls

software system

systems need to be “process aware”!

data centric  process centric
Process Aware Information Systems

Four types of "workflow-like" systems:

1. Information systems with **hard-coded** workflows (process& organization specific).

2. **Custom-made** information systems with **generic** workflow support (organization specific).

3. **Generic software with embedded workflow functionality** (e.g., the workflow components of ERP, CRM, PDM, etc. systems).

4. **Generic software focusing on workflow functionality** (e.g., Staffware, MQSeries Workflow, FLOWer, COSA, Oracle BPEL, Filenet, etc.).
The explosion of workflow systems in the mid 90-ties continues …

cf. Michael zur Mühlen
“Verification is important and feasible”

“The expressiveness of a PAIS system depends on the language used to configure the system”

“But, analysis of models only makes sense if they are an adequate reflection of reality”

Dual role of process models
- Specifies
- Configures
- Implements
- Analyzes
- Supports/controls

People, machines, components, organizations, business processes

Verification

Process/system model
Process mining: Linking events to models
Outline

1. Workflow patterns
2. Process verification
3. Process mining
Outline (2)

- Process mining
- Process verification
- Workflow patterns
- Diagnosis
- Process enactment
- Process design
- Implementation/configuration
Workflow Patterns
Workflow Patterns Initiative

• Started in 1999, joint work TU/e and QUT

• Objectives:
  – Identification of workflow modelling scenarios and solutions
  – Benchmarking
    • Workflow products (MQ/Series Workflow, Staffware, etc)
    • Proposed standards for web service composition (BPML, BPEL)
    • Process modelling languages (UML, BPMN)
  – Foundation for selecting workflow solutions

• Home Page: www.workflowpatterns.com

• Primary publication:

• Evaluations of commercial offerings, research prototypes, proposed standards for web service composition, etc
# The Workflow Patterns Framework

<table>
<thead>
<tr>
<th>Time</th>
<th>Control-flow P:s</th>
<th>Resource P:s</th>
<th>Data P:s</th>
<th>Revised Control-flow P:s</th>
</tr>
</thead>
<tbody>
<tr>
<td>2000</td>
<td>20</td>
<td>-</td>
<td>-</td>
<td>43</td>
</tr>
<tr>
<td>2003</td>
<td></td>
<td>-</td>
<td>-</td>
<td>43</td>
</tr>
<tr>
<td>Jun 2005</td>
<td></td>
<td>-</td>
<td>-</td>
<td>43</td>
</tr>
<tr>
<td>Oct 2005</td>
<td></td>
<td>-</td>
<td>-</td>
<td>43</td>
</tr>
<tr>
<td>Sep 2006</td>
<td></td>
<td>-</td>
<td>-</td>
<td>43</td>
</tr>
</tbody>
</table>

- Control-flow P:s 20
  - W. van der Aalst
  - A. ter Hofstede
  - B. Kiepuszewski
  - A. Barros

- Resource P:s - 43
  - N. Russell
  - W. van der Aalst
  - A. ter Hofstede
  - D. Edmond

- Data P:s - 40
  - N. Russell
  - W. van der Aalst
  - A. ter Hofstede
  - D. Edmond

- Revised Control-flow P:s 43
  - N. Russell
  - W. van der Aalst
  - A. ter Hofstede
  - N. Mulyar

- The ordering of activities in a process
  - CoopIS’2000
  - DAPD’2003

- Resource definition & work distribution in a process
  - CAiSE’2005

- Data representation and handling in a process
  - ER’2005

- 23 new patterns
- Formalised in CPN notation

---


[www.workflowpatterns.com](http://www.workflowpatterns.com)
The Workflow Patterns Framework

Evaluations

Control-flow P:s 20
- COSA
- FLOWer
- Eastman
- Meteor
- Mobile
- I-Flow
- Staffware
- InConcert
- Domino Workflow
- Visual Workflow
- Forte Conductor
- MQSeries/Workflow
- SAR R/3 Workflow
- Verve Workflow
- Changengine
- XPDL, BPEL4WS, BPML, WSFL, XLANG, WSCI, UML AD 1.4
- UML AD 2.0, BPMN

Resource P:s - 43
- Staffware
- WebSphere MQ
- FLOWer
- COSA
- iPlanet
- BPEL4WS
- UML AD 2.0
- BPMN

Data P:s - 40
- Staffware
- MQSeries
- FLOWer
- COSA
- iPlanet
- XPDL, BPEL4WS
- UML AD 2.0, BPMN

Language Development: YAWL/newYAWL
YAWL system
Process Verification

**Event Logs**

Models analyze events, e.g., messages, transactions, etc.

**Discovery**

Records events, e.g., messages, transactions, etc.

**Conformance**

Supports/controls

**Software System**

Specifies, configures, implements, analyzes

**World**

Business processes, people, machines, components, organizations

**Process/System Model**

Verifies, analyzes

**Event Logs**

OK/NOK
Example: Verification of the SAP Reference model (Joint work with Jan Mendling)

- The SAP reference model contains more than 600 non-trivial process models expressed in terms of Event-driven Process Chains (EPCs).
Approach

604 non-trivial process models

Collect characteristics

Model analysis

Compare
Simplistic approach: YAWL + invariants

Analysis using **transition invariants**, i.e., only lower bound! ProM allows for more precise analysis
Simplistic approach: YAWL + Petri net invariants
<table>
<thead>
<tr>
<th>Branch</th>
<th>Model</th>
<th>% EPC</th>
<th>% $E_{av.}$</th>
<th>$F_{av.}$</th>
<th>$C_{av.}$</th>
<th>$A_{av.}$</th>
<th>Cycle</th>
<th>Error</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asset Accounting</td>
<td>461</td>
<td>4.7%</td>
<td>43</td>
<td>7.1%</td>
<td>13.9</td>
<td>4.0</td>
<td>5.2</td>
<td>23.3</td>
<td>0</td>
</tr>
<tr>
<td>Benefits Administration</td>
<td>50</td>
<td>0.5%</td>
<td>6</td>
<td>1.0%</td>
<td>9.5</td>
<td>3.3</td>
<td>5.8</td>
<td>19.7</td>
<td>3</td>
</tr>
<tr>
<td>Compensation Management</td>
<td>122</td>
<td>1.2%</td>
<td>18</td>
<td>3.0%</td>
<td>7.6</td>
<td>3.4</td>
<td>3.3</td>
<td>13.7</td>
<td>3</td>
</tr>
<tr>
<td>Customer Service</td>
<td>402</td>
<td>4.1%</td>
<td>41</td>
<td>6.8%</td>
<td>16.5</td>
<td>3.6</td>
<td>9.0</td>
<td>29.5</td>
<td>3</td>
</tr>
<tr>
<td>Enterprise Controlling</td>
<td>599</td>
<td>6.1%</td>
<td>22</td>
<td>3.6%</td>
<td>14.3</td>
<td>10.1</td>
<td>6.1</td>
<td>32.1</td>
<td>0</td>
</tr>
<tr>
<td>Environment, Health, Safety</td>
<td>102</td>
<td>1.0%</td>
<td>19</td>
<td>3.1%</td>
<td>3.5</td>
<td>2.7</td>
<td>1.2</td>
<td>7.0</td>
<td>0</td>
</tr>
<tr>
<td>Financial Accounting</td>
<td>614</td>
<td>6.2%</td>
<td>54</td>
<td>8.9%</td>
<td>13.0</td>
<td>4.0</td>
<td>5.1</td>
<td>21.8</td>
<td>0</td>
</tr>
<tr>
<td>Position Management</td>
<td>4</td>
<td>0.0%</td>
<td>0</td>
<td>0.0%</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0</td>
</tr>
<tr>
<td>Inventory Management</td>
<td>184</td>
<td>1.9%</td>
<td>3</td>
<td>0.5%</td>
<td>15.0</td>
<td>7.0</td>
<td>6.0</td>
<td>28.0</td>
<td>2</td>
</tr>
<tr>
<td>Organizational Management</td>
<td>37</td>
<td>0.4%</td>
<td>5</td>
<td>0.8%</td>
<td>12.0</td>
<td>3.0</td>
<td>6.6</td>
<td>24.0</td>
<td>3</td>
</tr>
<tr>
<td>Payroll</td>
<td>541</td>
<td>5.5%</td>
<td>7</td>
<td>1.2%</td>
<td>5.7</td>
<td>3.1</td>
<td>2.1</td>
<td>11.4</td>
<td>0</td>
</tr>
<tr>
<td>Personnel Administration</td>
<td>15</td>
<td>0.2%</td>
<td>4</td>
<td>0.7%</td>
<td>7.3</td>
<td>1.5</td>
<td>4.0</td>
<td>12.3</td>
<td>0</td>
</tr>
<tr>
<td>Personnel Development</td>
<td>60</td>
<td>0.6%</td>
<td>10</td>
<td>1.7%</td>
<td>8.7</td>
<td>2.5</td>
<td>4.4</td>
<td>15.6</td>
<td>3</td>
</tr>
<tr>
<td>Personnel Time Management</td>
<td>87</td>
<td>0.9%</td>
<td>12</td>
<td>2.0%</td>
<td>10.8</td>
<td>3.0</td>
<td>5.3</td>
<td>19.5</td>
<td>1</td>
</tr>
<tr>
<td>Plant Maintenance</td>
<td>399</td>
<td>4.1%</td>
<td>35</td>
<td>5.8%</td>
<td>20.5</td>
<td>4.2</td>
<td>11.4</td>
<td>37.8</td>
<td>9</td>
</tr>
<tr>
<td>Procurement</td>
<td>444</td>
<td>4.5%</td>
<td>37</td>
<td>6.1%</td>
<td>6.7</td>
<td>3.5</td>
<td>2.7</td>
<td>12.4</td>
<td>0</td>
</tr>
<tr>
<td>Product Data Management</td>
<td>366</td>
<td>3.7%</td>
<td>26</td>
<td>4.3%</td>
<td>4.5</td>
<td>5.4</td>
<td>2.2</td>
<td>13.7</td>
<td>0</td>
</tr>
<tr>
<td>Production</td>
<td>206</td>
<td>3.0%</td>
<td>17</td>
<td>2.8%</td>
<td>8.8</td>
<td>3.0</td>
<td>2.9</td>
<td>13.7</td>
<td>0</td>
</tr>
<tr>
<td>Production Planning</td>
<td>194</td>
<td>2.0%</td>
<td>17</td>
<td>2.8%</td>
<td>5.7</td>
<td>2.9</td>
<td>3.0</td>
<td>11.5</td>
<td>0</td>
</tr>
<tr>
<td>Project Management</td>
<td>347</td>
<td>3.5%</td>
<td>36</td>
<td>6.0%</td>
<td>8.5</td>
<td>3.8</td>
<td>2.2</td>
<td>14.0</td>
<td>0</td>
</tr>
<tr>
<td>Quality Management</td>
<td>209</td>
<td>2.1%</td>
<td>20</td>
<td>3.3%</td>
<td>20.5</td>
<td>3.8</td>
<td>11.7</td>
<td>37.8</td>
<td>1</td>
</tr>
<tr>
<td>Real Estate Management</td>
<td>169</td>
<td>1.7%</td>
<td>6</td>
<td>1.0%</td>
<td>12.7</td>
<td>6.5</td>
<td>7.3</td>
<td>27.0</td>
<td>1</td>
</tr>
<tr>
<td>Recruitment</td>
<td>56</td>
<td>0.6%</td>
<td>9</td>
<td>1.5%</td>
<td>7.4</td>
<td>2.6</td>
<td>4.1</td>
<td>13.8</td>
<td>3</td>
</tr>
<tr>
<td>Retail</td>
<td>842</td>
<td>8.6%</td>
<td>1</td>
<td>0.2%</td>
<td>7.0</td>
<td>5.0</td>
<td>2.0</td>
<td>11.0</td>
<td>0</td>
</tr>
<tr>
<td>Revenue &amp; Cost Controlling</td>
<td>568</td>
<td>5.8%</td>
<td>19</td>
<td>3.1%</td>
<td>16.5</td>
<td>10.2</td>
<td>7.9</td>
<td>36.0</td>
<td>1</td>
</tr>
<tr>
<td>Sales &amp; Distribution</td>
<td>703</td>
<td>7.1%</td>
<td>76</td>
<td>12.6%</td>
<td>10.6</td>
<td>3.1</td>
<td>4.3</td>
<td>16.6</td>
<td>0</td>
</tr>
<tr>
<td>Training &amp; Event Management</td>
<td>95</td>
<td>1.0%</td>
<td>12</td>
<td>2.0%</td>
<td>13.0</td>
<td>2.7</td>
<td>6.2</td>
<td>22.2</td>
<td>0</td>
</tr>
<tr>
<td>Travel Management</td>
<td>116</td>
<td>1.2%</td>
<td>1</td>
<td>0.2%</td>
<td>24.0</td>
<td>7.0</td>
<td>16.0</td>
<td>48.0</td>
<td>0</td>
</tr>
<tr>
<td>Treasury</td>
<td>1761</td>
<td>17.9%</td>
<td>48</td>
<td>7.9%</td>
<td>10.5</td>
<td>3.5</td>
<td>4.5</td>
<td>18.1</td>
<td>0</td>
</tr>
<tr>
<td>All 29 Branches</td>
<td>9844</td>
<td>100%</td>
<td>604</td>
<td>100%</td>
<td>11.5</td>
<td>4.0</td>
<td>5.2</td>
<td>20.8</td>
<td>33</td>
</tr>
</tbody>
</table>
5.6% is a lower bound!

• Using more refined techniques more errors are found, e.g., using reduction rules and state-space analysis it can be shown that 20.9% of the SAP models are incorrect (126/604).

• Other large repositories of EPC models:
  – Collection of 381 non-trivial EPCs from a German process reengineering project in the service sector
  – Collection of 935 non-trivial EPCs from the Austrian financial industry
  – Collection of 83 non-trivial EPCs from three different consulting companies

• Total: 2003 non-trivial EPCs
Overview results

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Complete Sample</th>
<th>SAP Ref. Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>xOEPC errors</td>
<td>154</td>
<td>90</td>
</tr>
<tr>
<td>Unreduced EPCs</td>
<td>156</td>
<td>103</td>
</tr>
<tr>
<td>ProM error EPCs</td>
<td>115</td>
<td>75</td>
</tr>
<tr>
<td>EPCs with errors</td>
<td>215</td>
<td>126</td>
</tr>
<tr>
<td>EPCs in total</td>
<td>2003</td>
<td>604</td>
</tr>
<tr>
<td>Error ratio</td>
<td><strong>10.7%</strong></td>
<td><strong>20.9%</strong></td>
</tr>
</tbody>
</table>

- Designers make errors (10.7%)
- Errors can be predicted (95.2%)
- Process verification is mature, but models are not!
- Disconnect between ref. models and systems cf. SAP
Limitations of using models as a starting point

- Supports/controls
- Specifies/configures/implements/analyzes

"world" business processes people machines components organizations

verification

process/system model

real world

powerpoint reality
Process Mining

Process Mining software system process/system model event logs models analyzes discovery records events, e.g., messages, transactions, etc.

“world” business processes people machines components organizations

supports/controls specifies configures implements analyzes

discovery conformance

process/system model event logs

Petri net analysis Mining flexible processes

ProM - the leading process mining toolkit.

current version: 4.1 released 15.4.2007

control-flow process mining
Event logs are a reflection of reality

“logs are everywhere and there will be more …”
Examples:
Process mining: Linking events to models

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

models
- analyzes
- supports/controls
- specifies
- configures
- implements
- analyzes
- records
- events, e.g.,
- messages,
- transactions,
- etc.

process/system model

software system

event logs

verification

discovery

conformance
Toy example to explain basic idea:

Reviewing of papers for journal 😊
Event log:
• processes
  – process instances
• events

Per event:
• activity name
• (event type)
• (originator)
• (timestamp)
• (data)
<ProcessInstance id="51" description="">
  - <AuditTrailEntry>
    <WorkflowModelElement>invite reviewers</WorkflowModelElement>
    <EventType>start</EventType>
    <Timestamp>2006-08-28T00:00:00.000+01:00</Timestamp>
    <Originator>Mike</Originator>
  </AuditTrailEntry>
  - <AuditTrailEntry>
    <WorkflowModelElement>invite reviewers</WorkflowModelElement>
    <EventType>complete</EventType>
    <Timestamp>2006-08-31T00:00:00.000+01:00</Timestamp>
    <Originator>Mike</Originator>
  </AuditTrailEntry>
  - <AuditTrailEntry>
    <Data/>
    <Attribute name="result">reject</Attribute>
  </Data>
  <WorkflowModelElement>get review 3</WorkflowModelElement>
  <EventType>complete</EventType>
  <Timestamp>2006-09-02T00:00:00.000+01:00</Timestamp>
  <Originator>Mary</Originator>
  </AuditTrailEntry>
  - <AuditTrailEntry>
    <WorkflowModelElement>time out 1</WorkflowModelElement>
    <EventType>complete</EventType>
    <Timestamp>2006-09-03T00:00:00.000+01:00</Timestamp>
    <Originator />
  </AuditTrailEntry>
  - <AuditTrailEntry>
Discovery models analyzes discovery records events, e.g., messages, transactions, etc.
specifies configures implements analyzes supports/controls
conformance

people machines components organizations

process/system model

discovery

software system

event logs

verification

records events, e.g., messages, transactions, etc.
No transactional information
Corresponding EPC model (used by SAP, ARIS, etc)
YAWL model (executable workflow model)
about 30 mining plug-ins!
Social network analysis
Decision point analysis
Performance analysis
Discovering patterns
Conformance Checking

models analyzes discovers configures implements specifies supports/controls

process/system model event logs

verification discovery conformance

“world”

business processes
people machines components organizations records events, e.g., messages, transactions, etc.

software system
Comparing the discovered model with the log (f=1)
Different process model, same log ($f=0.796$)

Decision cannot be repeated according to model but can be repeated in reality!
Adding deviations to the log \((f=0.89)\)
LTL checker plug-in

hanging business out to dry.
Goal of ProM: Complete support

- Staffware
- FLOWer
- Websphere
- YAWL
- ADEPT
- ARIS PPM/SIM
- Outlook
- Caramba
- SAP
- PeopleSoft
- InConcert
- IBM MQSeries
- CPN Tools
- CVS
- Oracle BPEL
- UML SD
- company specific systems

SAP
PeopleSoft
InConcert
IBM MQSeries
CPN Tools
CVS
Oracle BPEL
UML SD
company specific systems

...
REALITY CHECK AHEAD
Reality Check

• Process mining on structured/administrative workflow-like logs is relatively easy.
• However, let us look at two *extreme* logs:
  – A log from a hospital with information on treatments, complications, and diagnoses.
  – A log from a manufacturer of high-tech system with information on system tests.
First example: Hospital data

- Information on treatment, complication, and diagnosis events.
- Data:
  - 2712 cases (all unique)
  - 29258 events
  - +/- 10.8 events per case
  - 264 different events (activities)
### Frequency of activities

<table>
<thead>
<tr>
<th>Model element</th>
<th>Event type</th>
<th>Occurrences (absolute)</th>
<th>Occurrences (relative)</th>
</tr>
</thead>
<tbody>
<tr>
<td>B_Perifeer infuus</td>
<td>start</td>
<td>2837</td>
<td>9.696%</td>
</tr>
<tr>
<td>B_Maagsonde</td>
<td>start</td>
<td>2430</td>
<td>8.305%</td>
</tr>
<tr>
<td>B_Beademing</td>
<td>start</td>
<td>2187</td>
<td>7.475%</td>
</tr>
<tr>
<td>B_Catheter a Demeure</td>
<td>start</td>
<td>2096</td>
<td>7.164%</td>
</tr>
<tr>
<td>B_Basiszorg</td>
<td>start</td>
<td>2010</td>
<td>6.87%</td>
</tr>
<tr>
<td>B_Arterie lijn op OK</td>
<td>start</td>
<td>2002</td>
<td>6.843%</td>
</tr>
<tr>
<td>B_O2 masker/slang</td>
<td>start</td>
<td>1954</td>
<td>6.679%</td>
</tr>
<tr>
<td>B_Thoraxdrain</td>
<td>start</td>
<td>1863</td>
<td>6.367%</td>
</tr>
<tr>
<td>Code</td>
<td>Beschrijving</td>
<td>Start</td>
<td>Frequentie</td>
</tr>
<tr>
<td>------</td>
<td>--------------------------------------------------</td>
<td>-------</td>
<td>------------</td>
</tr>
<tr>
<td>C_N</td>
<td>Phrenicus Paralyse</td>
<td>1</td>
<td>0,003%</td>
</tr>
<tr>
<td>C_TIA</td>
<td>TIA</td>
<td>1</td>
<td>0,003%</td>
</tr>
<tr>
<td>B_Horizontaal</td>
<td></td>
<td>1</td>
<td>0,003%</td>
</tr>
<tr>
<td>C_Cholecystitis, acalc</td>
<td></td>
<td>1</td>
<td>0,003%</td>
</tr>
<tr>
<td>C_Decubitus hak st. 3a</td>
<td></td>
<td>1</td>
<td>0,003%</td>
</tr>
<tr>
<td>C_Druk necrose elders</td>
<td></td>
<td>1</td>
<td>0,003%</td>
</tr>
<tr>
<td>B_Decubitus zorg stadium 3b</td>
<td></td>
<td>1</td>
<td>0,003%</td>
</tr>
<tr>
<td>C_Haemolyse</td>
<td></td>
<td>1</td>
<td>0,003%</td>
</tr>
<tr>
<td>B_Decubitus zorg stadium 4b</td>
<td></td>
<td>1</td>
<td>0,003%</td>
</tr>
<tr>
<td>B_Isolatie Beschermend</td>
<td></td>
<td>1</td>
<td>0,003%</td>
</tr>
<tr>
<td>B_Donor Weefsel</td>
<td></td>
<td>1</td>
<td>0,003%</td>
</tr>
<tr>
<td>C_Polyurie (&gt;40ml/kg/24u)</td>
<td></td>
<td>1</td>
<td>0,003%</td>
</tr>
<tr>
<td>C_Decubitus overig st. 3a</td>
<td></td>
<td>1</td>
<td>0,003%</td>
</tr>
<tr>
<td>C_Intra-peritoneaal Abces</td>
<td></td>
<td>1</td>
<td>0,003%</td>
</tr>
</tbody>
</table>
Selection: Care after heart surgery

• Data
  – 874 cases (all unique)
  – 10478 events
  – 181 different events (activities)
Second example: Test data from high-tech system manufacturer

- Information on testing process of high-tech systems.
- Data:
  - 24 comparable cases
  - 154966 events
  - +/- 6450 events per case
  - between 2820 and 16250 events per machine
  - 720 different events (start/complete activities)
Helicopter view
Average time spent in job-steps (aggregated events)
Mining just the complete events (# 360)…
Common activities (#70)
Job step level
Conformance checker (reference model – job steps)
Discovered models fit better than reference model

<table>
<thead>
<tr>
<th>Machine ID</th>
<th>Fitness with respect to the reference process model</th>
<th>Fitness with respect to the discovered process model</th>
</tr>
</thead>
<tbody>
<tr>
<td>0431</td>
<td>$f = 0.30895045$</td>
<td>$f = 0.75113416$</td>
</tr>
<tr>
<td>0278</td>
<td>$f = 0.38491702$</td>
<td>$f = 0.82790506$</td>
</tr>
<tr>
<td>0185</td>
<td>$f = 0.37574032$</td>
<td>$f = 0.7171429$</td>
</tr>
<tr>
<td>0466</td>
<td>$f = 0.35643995$</td>
<td>$f = 0.74496365$</td>
</tr>
<tr>
<td>0391</td>
<td>$f = 0.38410193$</td>
<td>$f = 0.72710335$</td>
</tr>
<tr>
<td>1722</td>
<td>$f = 0.33359805$</td>
<td>$f = 0.7599035$</td>
</tr>
<tr>
<td>1694</td>
<td>$f = 0.39662793$</td>
<td>$f = 0.7821204$</td>
</tr>
<tr>
<td>1256</td>
<td>$f = 0.40988705$</td>
<td>$f = 0.7436544$</td>
</tr>
<tr>
<td>1343</td>
<td>$f = 0.3985147$</td>
<td>$f = 0.70074475$</td>
</tr>
<tr>
<td>1981</td>
<td>$f = 0.35723096$</td>
<td>$f = 0.6668054$</td>
</tr>
<tr>
<td>1754</td>
<td>$f = 0.401513$</td>
<td>$f = 0.77557445$</td>
</tr>
<tr>
<td>1662</td>
<td>$f = 0.4138763$</td>
<td>$f = 0.76897943$</td>
</tr>
<tr>
<td>1453</td>
<td>$f = 0.40458512$</td>
<td>$f = 0.5956065$</td>
</tr>
<tr>
<td>1298</td>
<td>$f = 0.37758428$</td>
<td>$f = 0.42357332$</td>
</tr>
<tr>
<td>1876</td>
<td>$f = 0.35556892$</td>
<td>$f = 0.7529762$</td>
</tr>
<tr>
<td>1656</td>
<td>$f = 0.36802232$</td>
<td>$f = 0.65629774$</td>
</tr>
<tr>
<td>1099</td>
<td>$f = 0.424476$</td>
<td>$f = 0.67167395$</td>
</tr>
<tr>
<td>1919</td>
<td>$f = 0.33690846$</td>
<td>$f = 0.72738254$</td>
</tr>
<tr>
<td>1348</td>
<td>$f = 0.41031277$</td>
<td>$f = 0.63753587$</td>
</tr>
<tr>
<td>1596</td>
<td>$f = 0.41015995$</td>
<td>$f = 0.58138484$</td>
</tr>
<tr>
<td>1164</td>
<td>$f = 0.37603533$</td>
<td>$f = 0.67173433$</td>
</tr>
<tr>
<td>1032</td>
<td>$f = 0.32361075$</td>
<td>$f = 0.7062931$</td>
</tr>
<tr>
<td>1794</td>
<td>$f = 0.39387232$</td>
<td>$f = 0.7337234$</td>
</tr>
<tr>
<td>1160</td>
<td>$f = 0.40484217$</td>
<td>$f = 0.7697767$</td>
</tr>
</tbody>
</table>
Research challenge

*Mining less structured processes: the more unstructured, the more important it is to know what is going on!*
More significant nodes are emphasized

Highlights more important paths
More to learn from maps...

**Aggregation**
Clustering of coherent, less significant structures

**Abstraction**
Removing isolated, less significant structures
ProM’s Frequency abstraction miner
Conclusion (1)

Process verification
Process mining
Workflow patterns

process enactment
implementation/configuration

diagnosis
process design
Conclusion (2)

- Reality is different from models!
- The existence of event data enables a wide variety of process mining techniques: discovery and conformance.
- ProM supports this (+150 plug-ins)
- Although quite successful for "structured processes", "spaghetti processes" remain a challenge (two examples were given).
- Research should aim to address this challenge.
Relevant WWW sites

- http://www.processmining.org
- http://www.workflowpatterns.com
- http://www.workflowcourse.com
- http://www.win.tue.nl/is/
- http://is.tm.tue.nl/staff/wvdaalst