Process Mining

- Process discovery: "What is really happening?"
- Conformance checking: "Do we do what was agreed upon?"
- Performance analysis: "Where are the bottlenecks?"
- Process prediction: "Will this case be late?"
- Process improvement: "How to redesign this process?"
- Etc.
• Process discovery: "What is the real curriculum?"
• Conformance checking: "Do students meet the prerequisites?"
• Performance analysis: "Where are the bottlenecks?"
• Process prediction: "Will a student complete his studies (in time)?"
• Process improvement: "How to redesign the curriculum?"
Outline

• Trends in BPM
• Process Mining: The Basics
  • Input data
  • Discovery
  • Conformance
  • Software support
• Process Mining: Applications
• Process Mining: TomTom for Business Processes
• Conclusion
Trends in BPM
The first workflow management systems (called "office automation systems") were implemented in seventies, cf. Petri-net-based systems such as Officetalk (Xerox Parc, Skip Ellis) and SCOOP (Wharton, Michael Zisman).

Mid nineties: "explosion" of workflow products.

Shift from workflow automation to business process management.
Workflow Patterns Initiative

• Initiative started in late 90-ties.
• Collections:
  • 43 control-flow patterns (process/routing)
  • 40 data patterns
  • 43 resource patterns (work distr. and organization)
  • exception, flexibility, service interaction, ... patterns
• Frequently used as a tool in selection processes.
• Influenced standards (BPMN, BPEL, etc.) and systems.
• See www.workflowpatterns.com (+/- 500 unique visitors per day)
Problem is NOT the automation of structured processes!

Alignment
(Avoiding PowerPoint reality)

Ensuring compliance

Supporting flexibility
Where to start?

- process control
- process enactment
- implementation/configuration
- process design
- diagnosis

(arrow diagram connecting all terms)
Process Mining: The Basics
Role of models

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

supports/controls

software system

analyzes

specifies configures implements

"real world"

"powerpoint reality"

process/system model
Event logs are a reflection of reality
Process mining: Linking events to models
Starting point: event logs

event logs, audit trails, databases, message logs, etc.

unified event log (MXML)
What to discover?

• process models (Petri nets, EPCs, BPMN, etc.),
• organizational models,
• social networks,
• sequence diagrams,
• business rules,
• bottlenecks,
• simulation models,
• etc.

i.e., beyond "slice and dice" and showing KPIs on a dashboard ...
MXML Log
- instances: 3512
- audit trail entries: 46138

ProM supports +40 types of model discovery!
bottlenecks

flow time from A to B

throughput time
46138 events
A bit of theory:
Process discovery techniques

- Algorithmic techniques
  - Alpha miner
  - Alpha+, Alpha++, Alpha#
  - Heuristic miner
  - Multi phase miner
  - ...
- Genetic process mining
- Region-based process mining
  - State-based regions
  - Language based regions

cf. www.processmining.org for an overview
Example: Genetic Mining

1. initial population
2. fitness test
3. select best parents
4. crossover
5. children
6. mutation
7. new population

used in e.g. ProM, Futura Reflex, BPM|one
Conformance Checking

“world”
business processes
people
machines
components
organizations
models
analyzes

supports/
controls

software
system

specifies
configures
implements
analyzes

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

process/
system
model

discovery

conformance

event
logs
Conformance Checking

- Compare process model and event log: highlight deviations and measure conformance.
- Compare constraints/business rules and event logs: check e.g. the 4-eyes principle.
Tool support
• Open source initiative started in 2003 after several early prototypes.
• Common Public License (CPL).
• Current version: 5.0.
• ProMimport: to extract MXML from all kinds of applications
• Plug-in architecture.
• About 250 plug-ins available:
  • mining plug-ins: 38 (all mining algorithms presented and many more)
  • analysis plug-ins: 71 (e.g., verification, SNA, LTL, conformance checking, etc.)
  • import: 21 (for loading EPCs, Petri nets, YAWL, BPMN, etc.)
  • export: 44 (for storing EPCs, Petri nets, YAWL, BPMN, BPEL, etc.)
  • conversion: 45 (e.g., translating EPCs or BPMN into Petri nets)
  • filter: 24 (e.g., removing infrequent activities)
Screenshot of ProM 5.0
Business Intelligence Tools?

- Business Objects (SAP)
- Cognos Business Intelligence (IBM)
- Oracle Business Intelligence
- Hyperion (Oracle)
- SAS Business Intelligence
- Microsoft Business Intelligence
- SAP Business Intelligence (SAP BI)
- Jaspersoft (Open Source Business Intelligence)
- Pentaho BI Suite (Open Source)
- ....

- Dashboards, reports, scorecards, ...
- Slicing and dicing, data mining, ...
Process Mining Software

Futura Reflect

Comprehend

Pallas Athena

BPM|one

Interstage Automated Business Process Discovery & Visualization

ARIS Process Performance Manager

FUTURO

IDS Scheer

Process Discovery Focus

Business & Technology Optimization

Enterprise Visualization Suite
Process Mining: Applications
Where did we apply process mining?

- Municipalities (e.g., Alkmaar, Heusden, Harderwijk, etc.)
- Government agencies (e.g., Rijkswaterstaat, Centraal Justitieel Incasso Bureau, Justice department)
- Insurance related agencies (e.g., UWV)
- Banks (e.g., ING Bank)
- Hospitals (e.g., AMC hospital, Catharina hospital)
- Multinationals (e.g., DSM, Deloitte)
- High-tech system manufacturers and their customers (e.g., Philips Healthcare, ASML, Thales)
- Media companies (e.g. Winkwaves)
- ...
Example: A Dutch Municipality

144 cases
1326 events
Conformance check of discovered model

- Conformance check is sometimes not performed.
- Good fit is 97.9%.
- Drill down performed while not allowed.
- Activity is sometimes not performed.
- Both
Performance analysis
Events sorted by start time of case
Events sorted by duration
Idle time versus working time
"Real" animation
And of course ...
Reality ≠ PowerPoint (or Visio)
Process spectrum

structured
(Lasagna)

unstructured
(Spaghetti)
375 houses
18640 events
82 different activities
2712 patients
29258 events
264 different activities
874 patients
10478 events
181 different activities
24 machines
1549666 events
360 different activities
37.5% OK
62.5% NOK

design

reality
Process Mining: TomTom for Business Processes

TomTom

TU/e

Technische Universiteit Eindhoven
University of Technology

Where innovation starts
Business Process Navigation?

- Often a good process map is missing (incorrect, outdated, no color, ...)
- Process maps inherit the limitations of paper maps (no zoom or views)
- Process maps tend to aim at "controlling the driver"
- Current location unknown
- No traffic information is given
- No recalculation of the route
- No estimated arrival time
- ...
What we can learn from maps ...
Why imitate paper maps?

• Zoom in - zoom out
• Various views (e.g. show hotels and fuel stations at will)
• Dynamic content!
• Traffic information
• Show current location
ProM's Fuzzy Miner: Seamless zoom
ProM's "real animation"
ProM's "real simulation"
Prediction and recommendation

• Prediction: When are we home?
• Recommendation: What should I do next?
• Suggestions without force and the willingness to continuously recalculate the route.
ProM's Case prediction capabilities

144 cases
1326 events
Conclusion
Conclusion

• The abundance of event data enables a wide variety of process mining techniques ranging from process discovery to conformance checking.
• This is already possible today!
• Check out ProM with its 250+ plug-ins.
• A reality check for people that are involved in process modeling.
• Demand TomTom functionality!
Thanks! cf. www.processmining.org

- Wil van der Aalst
- Peter van den Brand
- Boudewijn van Dongen
- Christian Günther
- Eric Verbeek
- Ana Karla Alves de Medeiros
- Anne Rozinat
- Minseok Song
- Ton Weijters
- Remco Dijkman
- Gianluigi Greco
- Antonella Guzzo
- Kristian Bisgaard Lassen
- Ronny Mans
- Jan Mendling
- Vladimir Rubin
- Kenny van Uden
- Irene Vanderfeesten
- Barbara Weber
- Lijie Wen
- Mercy Amiyo
- Carmen Bratosin
- Toon Calders
- Jorge Cardoso
- Ronald Crooy
- Florian Gottschalk
- Monique Jansen-Vullers
- Peter Khisa Wakholi
- Nicolas Knaak
- Sven Lambrechts
- Joyce Nakatumba
- Mariska Netjes
- Mykola Pechenizkiy
- Maja Pesic
- Hajo Reijers
- Stefanie Rinderle
- Domenico Saccà
- Helen Schonenberg
- Marc Voorhoeve
- Jianmin Wang
- Jan Martijn van der Werf
- Martin van Wingerden
- Jianhong Ye
- Huub de Beer
- Elena Casares
- Alina Chipaila
- Walid Gaaloul
- Martijn van Giessel
- Shaifali Gupta
- Thomas Hoffmann
- Peter Hornix
- René Kerstjens
- Ralf Kramer
- Wouter Kunst
- Laura Maruster
- Andriy Nikolov
- Adarsh Ramesh
- Jo Theunissen
- ...
Relevant WWW sites

- http://www.processmining.org
- http://www.workflowpatterns.com
- http://www.workflowcourse.com
- http://www.vdaalst.com