ProM 4.0

Comprehensive Support for Real Process Analysis

Wil van der Aalst, Boudewijn van Dongen, Christian Günther, Ronny Mans, Ana Karla Alves de Medeiros, Anne Rozinat, Vladimir Rubin, Minseok Song, Eric Verbeek, and Ton Weijters

Eindhoven University of Technology,
P.O. Box 513, 5600 MB Eindhoven, The Netherlands
w.m.p.v.d.aalst@tue.nl

Motivation:
Logs are everywhere!

staffware. THE POWER OF PROCESS

IBM

océ

FileNet An IBM Company

Pallas Athena PROCESS MANAGEMENT

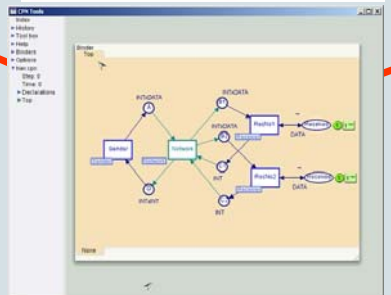
ORACLE

SAP

PHILIPS sense and simplicity

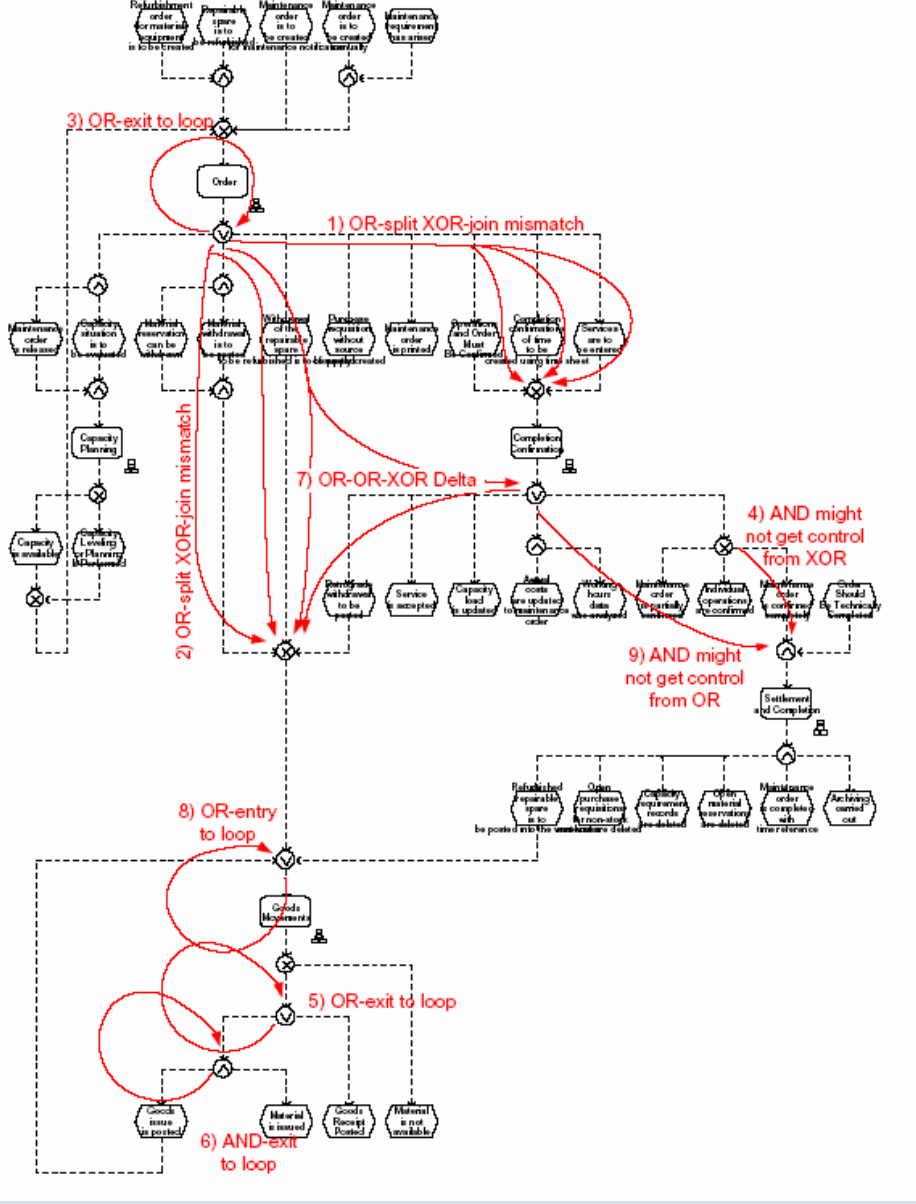
ASML

WebSphere. software



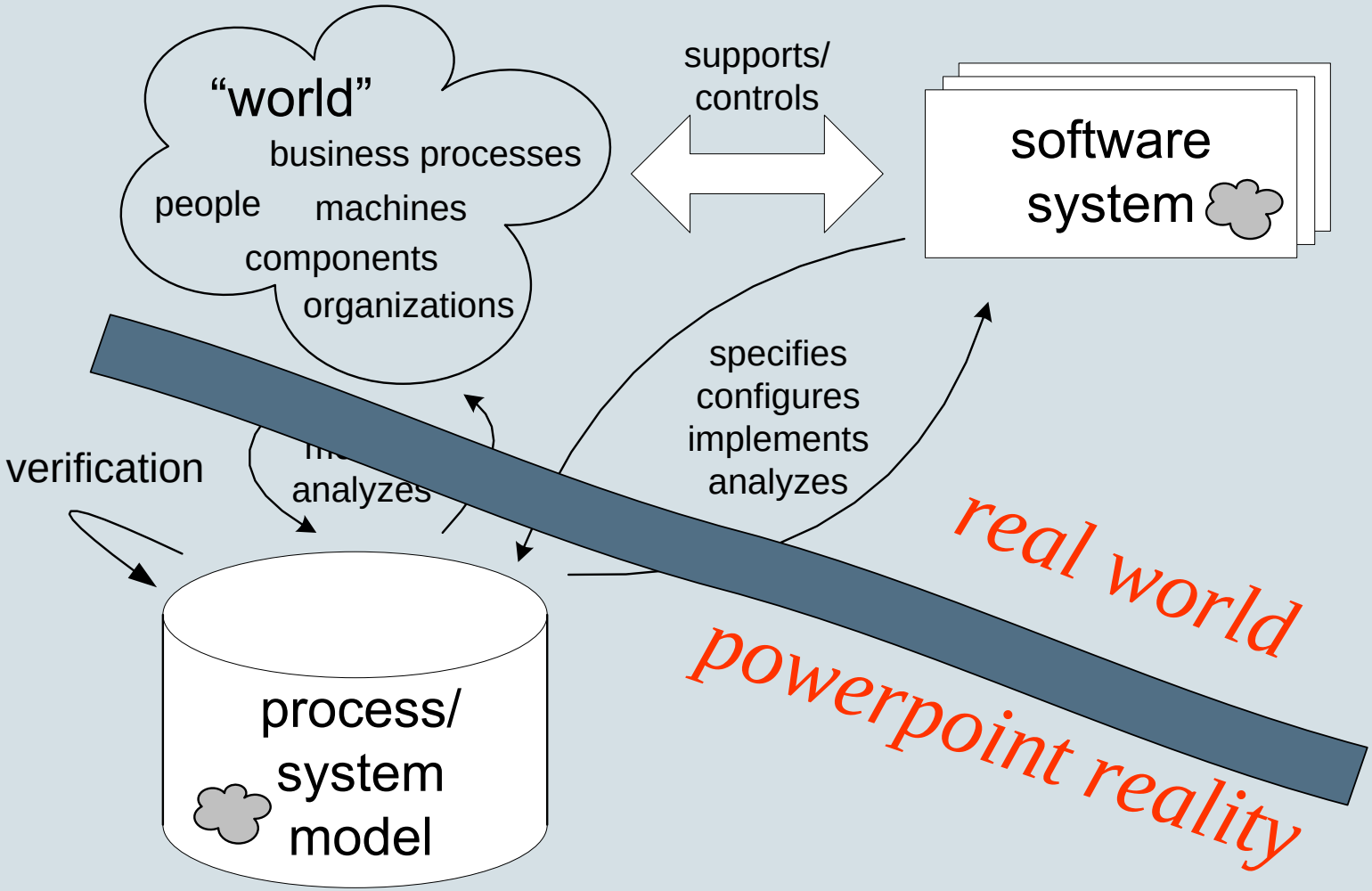
Motivation:
Man-made models typically suck!

Parameter Parameter	Complete Sample	SAP Ref. Model
xoEPC errors	154	90
Unreduced EPCs	156	103
ProM error EPCs	115	75
EPCs with errors	215	126
EPCs in total	2003	604
Error ratio	10.7%	20.9%

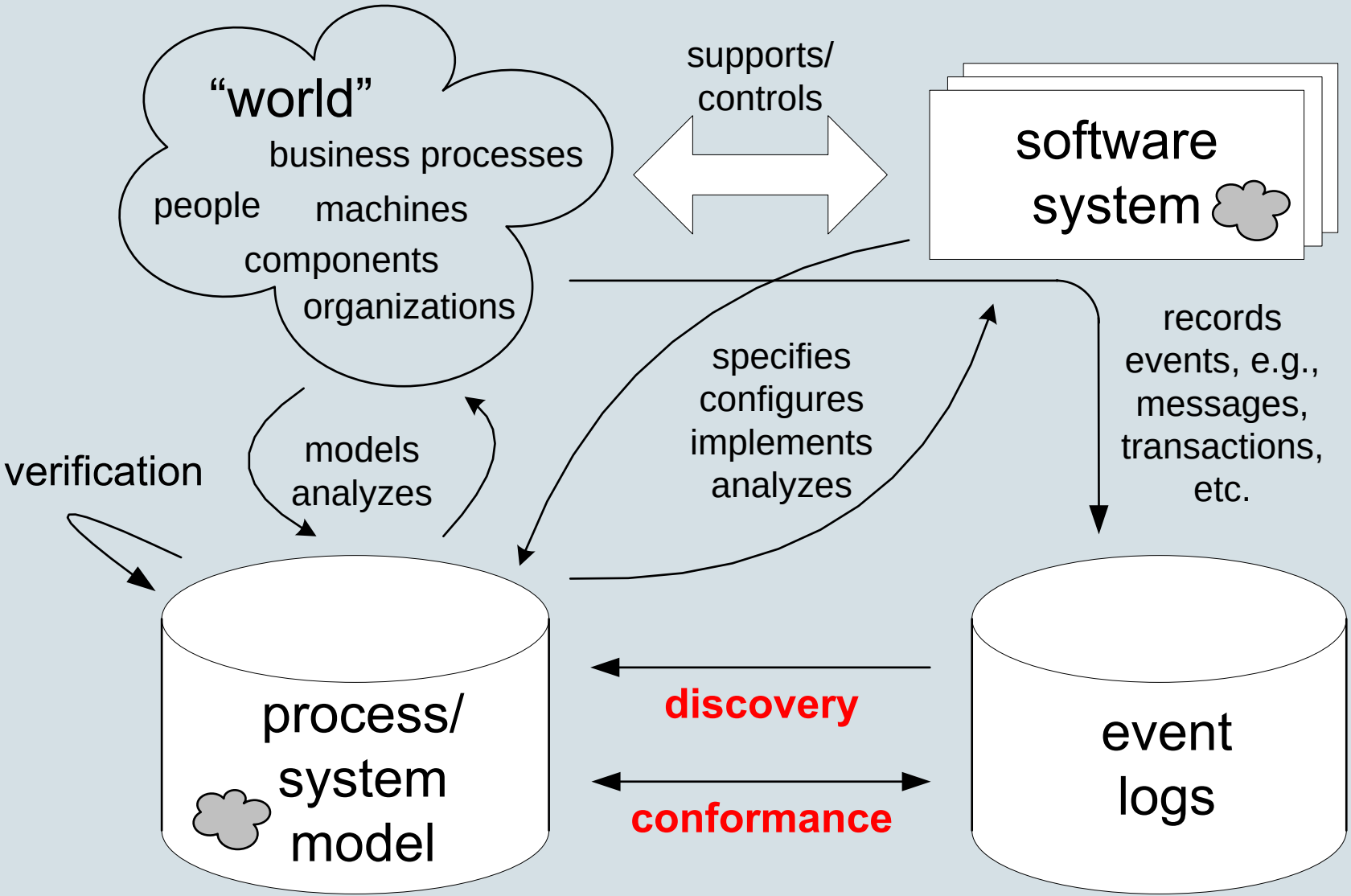


- Designers make errors (10.7
- Errors can be predicted (95.
- Process verification is mature, but models are not!
- Disconnect between ref. models and systems cf. SAP

Good models are needed, yet ...



Process mining: Linking events to models

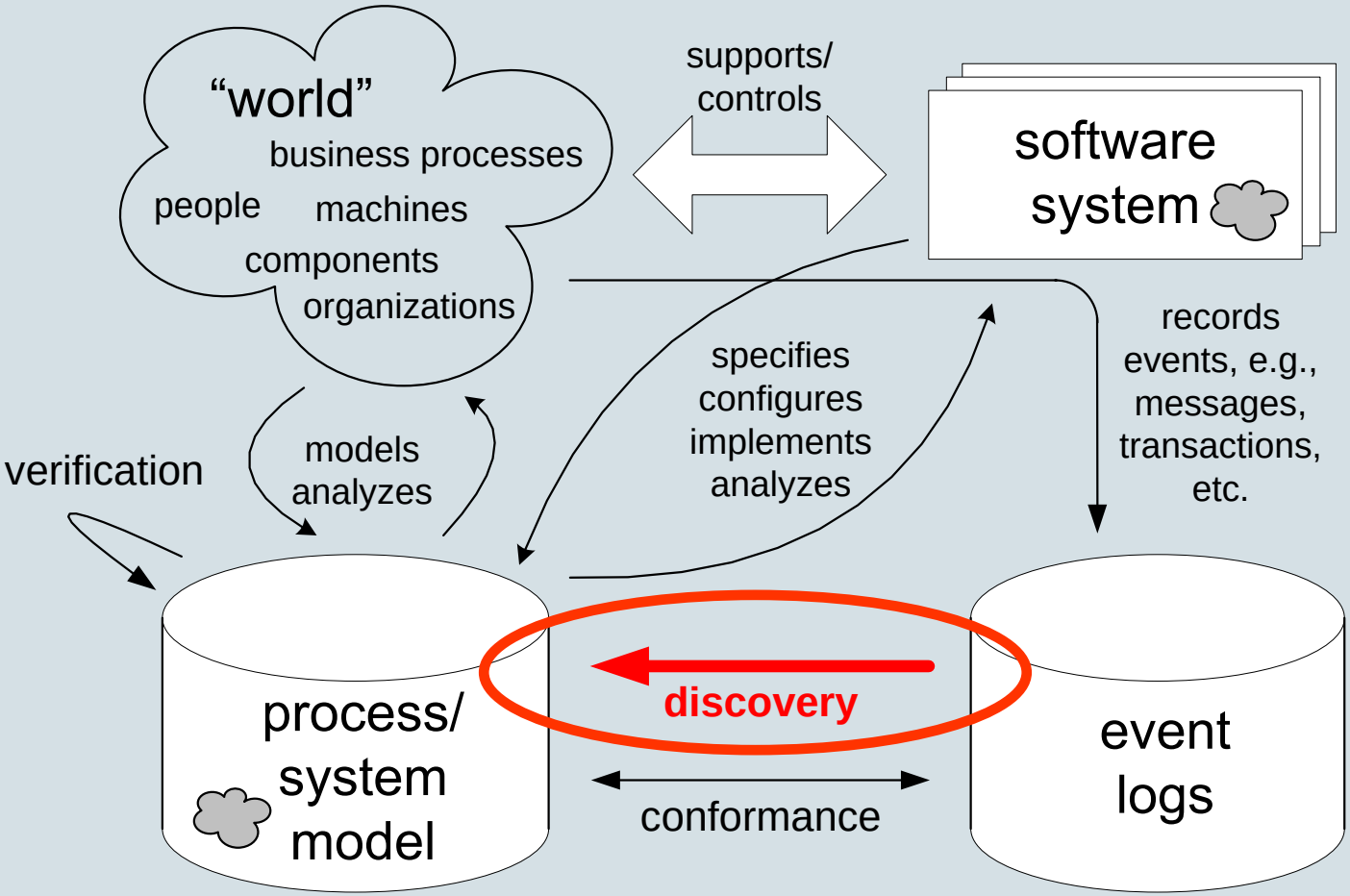


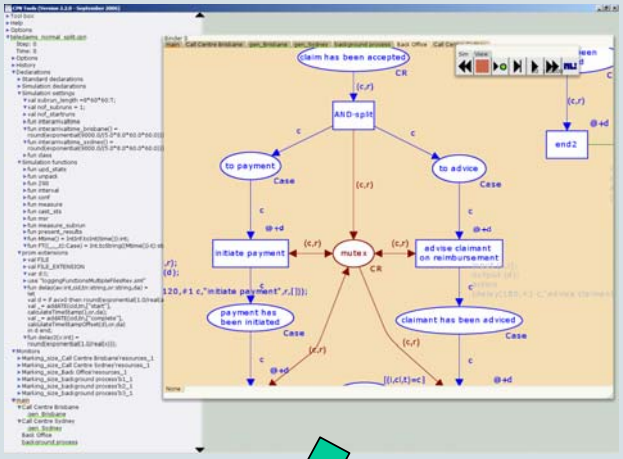
ProM

- First serious release: ProM 1.1 (2005)
- Paper: ProM 4.0
- Current release: ProM 4.1 (157 plug-ins)
- Nightly Builds (>185 plug-ins)

Version	ProM 1.1	ProM 4.0
Mining plug-ins	6	27
Analysis plug-ins	7	35
Import plug-ins	4	16
Export plug-ins	9	28
Conversion plug-ins	3	22
Log filter plug-ins	0	14
Total number of plug-ins	29	142

Discovery


"world"
business processes
people
machines
components
organizations
supports/
controls
software
system
records
events, e.g.,
messages,
transactions,
etc.
specifies
configures
implements
analyzes
models
analyzes
verification
process/
system
model
discovery
conformance
event
logs



MXML Log
- instances: **3512**
- audit trail entries: **46138**

- Control-flow Discovery
  - heuristics miner
  - genetic miner
  - multi-phase miner
  - region-based miner
  - fuzzy miner
  - ...
- Social network miner
- Performance miner
- Decision point miner
- ...

>40 plug-ins support some form of model discovery!

Heuristics miner

robust, dealing with noise, …

ProM [4.1]

File Mining Analysis Conversion Exports Window Help

Results - Settings for mining Filtered cpnToolsSimulationLog.mxml using Heuristics miner

incoming claim (complete) 3512

ProM [4.1]

File Mining Analysis Conversion Exports Window Help

Conversion - Heuristic net to Petri net

incoming claim complete

B check if sufficient information is available complete

S check if sufficient information is available complete

B register claim complete

S register claim complete

determine likelihood of claim complete

assess claim complete

ProM [4.1]

File Mining Analysis Conversion Exports Window Help

ProM

Conversion - Heuristic net to EPC

Status change to S register claim complete

ProM [4.1]

File Mining Analysis Conversion Exports Window Help

ProM

Conversion - Labeled WF net to YAWL model

[default]

incoming claim complete Task

[default]

true [1]

B check if sufficient information is available complete Task

true [0]

B register claim complete Task

S check if sufficient information is available complete Task

[default]

S register claim complete Task

determine likelihood of claim complete Task

true [0]

assess claim complete Task

true [0]

[default]

advise claimant on reimbursement complete Task

initiate payment complete Task

close claim complete Task

true [0]

end complete Task

[default]

true [0]

3512

Genetic miner

dealing with noise, time-consuming, …

ProM [4.1]

File Mining Analysis Conversion Exports Window Help

Settings for mining Filtered cpnToolsSimulationLog.mxml using Gene

Population size	100
Initial population type	Possible Duplicates
Max number generations	1000
Seed	1
Power value	1
Elitism rate	0.02
Fitness type	ExtraBehaviorPunishment

Show Advanced Fitness Parameters

Use genetic operators

Selection method type	Tournament 5
Crossover type	Enhanced
Crossover rate	0.8
Mutation type	Enhanced
Mutation rate	0.2

Help

ProM [4.1]

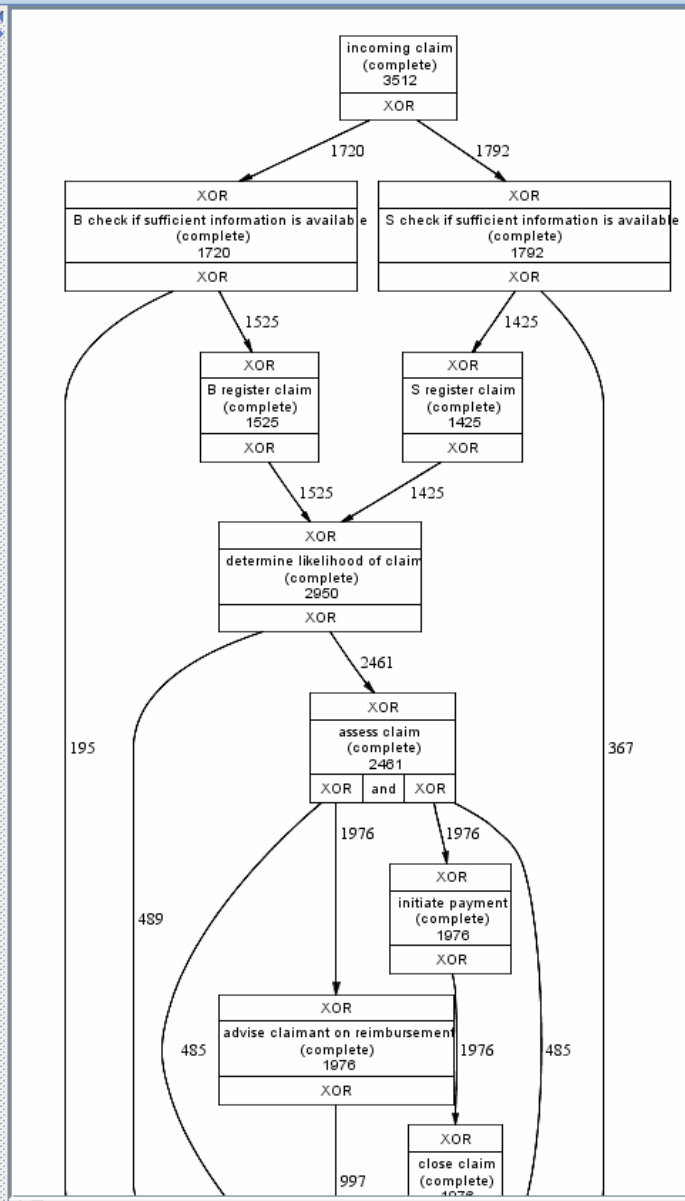
File Mining Analysis Conversion Exports Window Help

ProM

Results - Settings for mining Filtered cpnToolsSimulationLog.mxml using Genetic algorithm plugin

Population

Individual 0, fitness 0.98068908145235...
Individual 1, fitness 0.98018773937692...
Individual 2, fitness 0.98018213169065...
Individual 3, fitness 0.980106468971773
Individual 4, fitness 0.98008789790969...
Individual 5, fitness 0.97953950576728...
Individual 6, fitness 0.97949332608838...
Individual 7, fitness 0.97942139663262...
Individual 8, fitness 0.97902882274246...
Individual 9, fitness 0.97901433665930...
Individual 10, fitness 0.9787614169501...
Individual 11, fitness 0.9785506341884...
Individual 12, fitness 0.9784594547199...
Individual 13, fitness 0.9780242014563...
Individual 14, fitness 0.9779380159460...
Individual 15, fitness 0.9776391813049...
Individual 16, fitness 0.9770556016569...
Individual 17, fitness 0.97669507657444
Individual 18, fitness 0.9751499333233...
Individual 19, fitness 0.9736221193551...
Individual 20, fitness 0.9689060899081...
Individual 21, fitness 0.9667094325380...
Individual 22, fitness 0.9655598568529...
Individual 23, fitness 0.9640457178704...
Individual 24, fitness 0.9623202787780...
Individual 25, fitness 0.9600941185612...
Individual 26, fitness 0.9466830842117...
Individual 27, fitness 0.9225693557543...
Individual 28, fitness 0.9219690525070...
Individual 29, fitness 0.9215628029305...
Individual 30, fitness 0.9205037963208...
Individual 31, fitness 0.9198543398515...
Individual 32, fitness 0.9181237498297...
Individual 33, fitness 0.9175673186094...
Individual 34, fitness 0.9172804445404...
Individual 35, fitness 0.9171042198486...
Individual 36, fitness 0.9086432137727...
Individual 37, fitness 0.9080840831349...
Individual 38, fitness 0.9080784754486...
Individual 39, fitness 0.9064652001106...
Individual 40, fitness 0.9064588240428...
Individual 41, fitness 0.9058581346008...
Individual 42, fitness 0.9053089426693...
Individual 43, fitness 0.9024228757977...
Individual 44, fitness 0.9023710884326...
Individual 45, fitness 0.9021761030638...
Individual 46, fitness 0.9021439771387...
Individual 47, fitness 0.9010327666058...
Individual 48, fitness 0.8998150872721...
Individual 49, fitness 0.8905122495031...
Individual 50, fitness 0.8837302323569...
Individual 51, fitness 0.8630247459438...



Zoom: 82 %

Update graph

Display split/join semantics

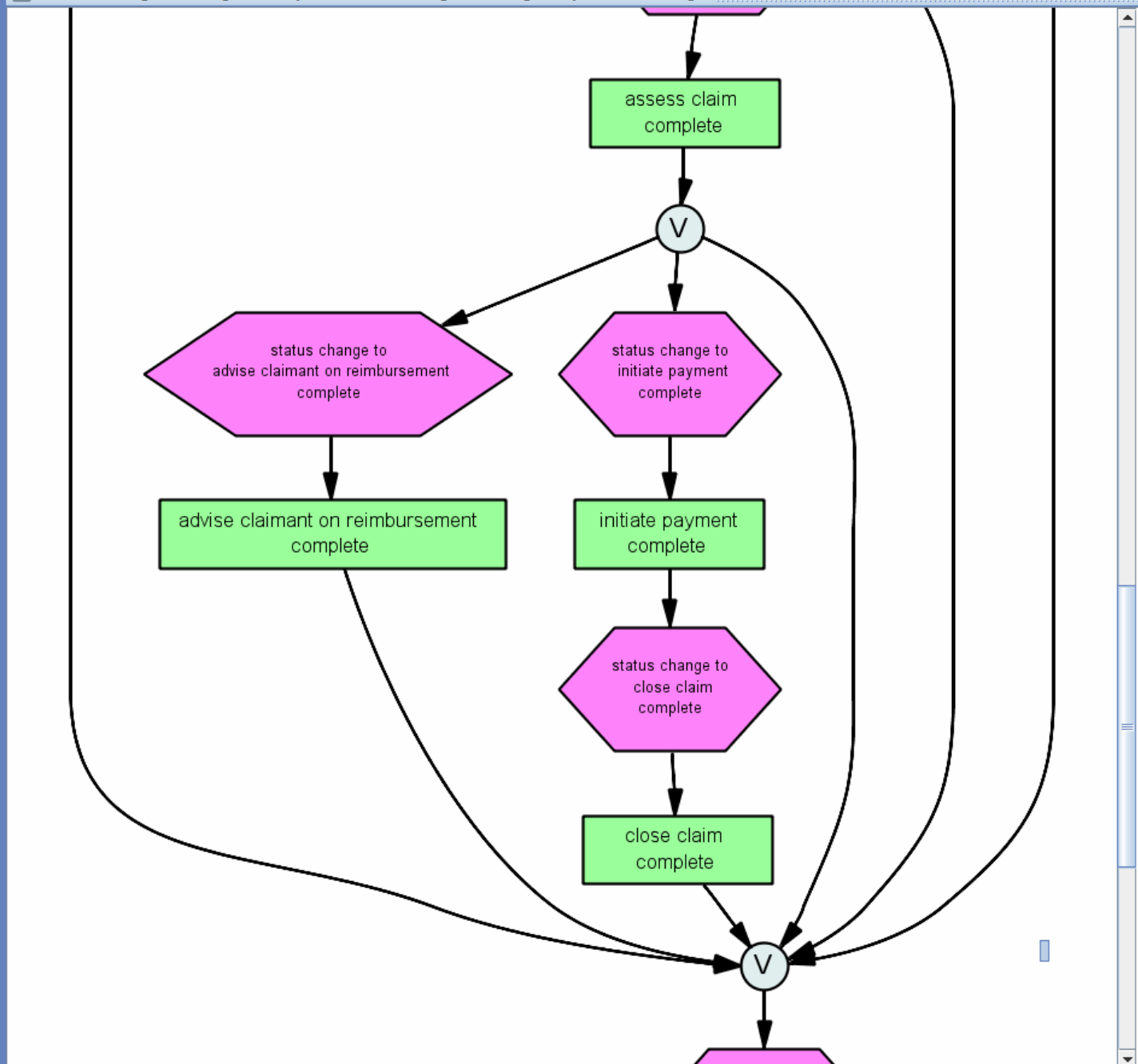
Multi-phase miner

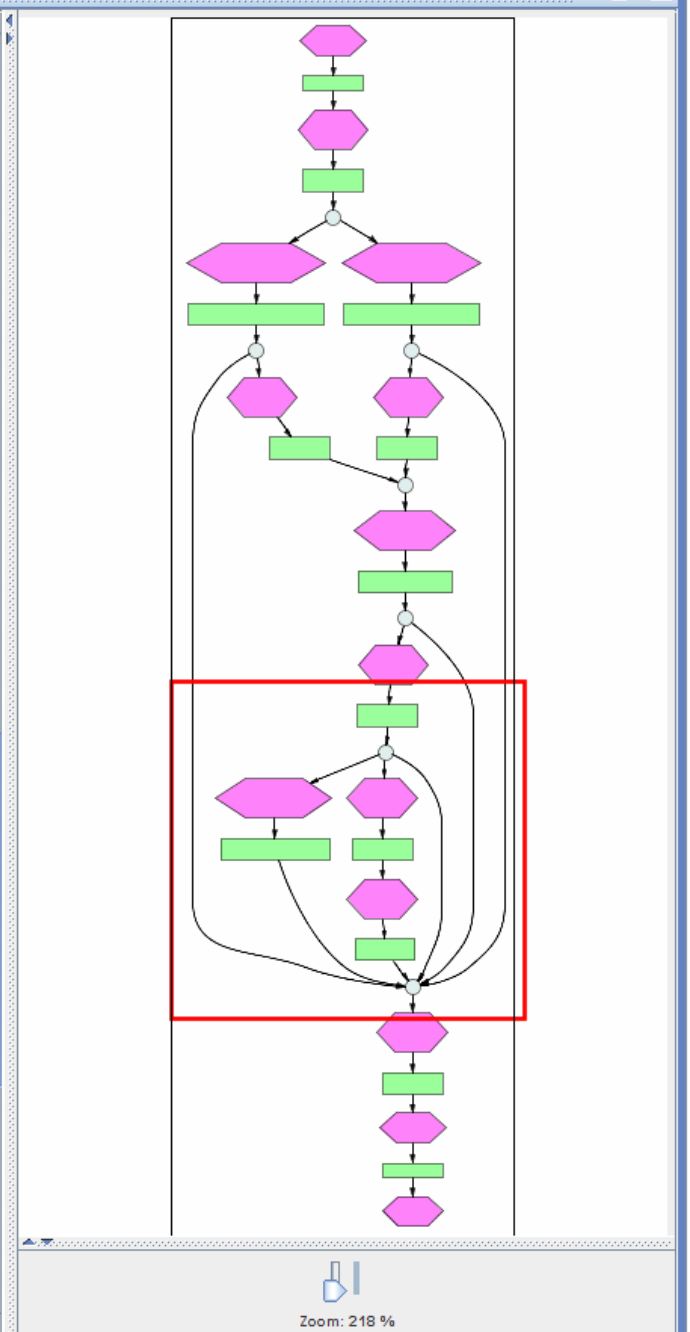
always a fitting model, over-generalize, …

ProM [4.1]

File Mining Analysis Conversion Exports Window Help

Results - Settings for mining Filtered cpnToolsSimulationLog.mxml using Multi-phase Macro Plugin

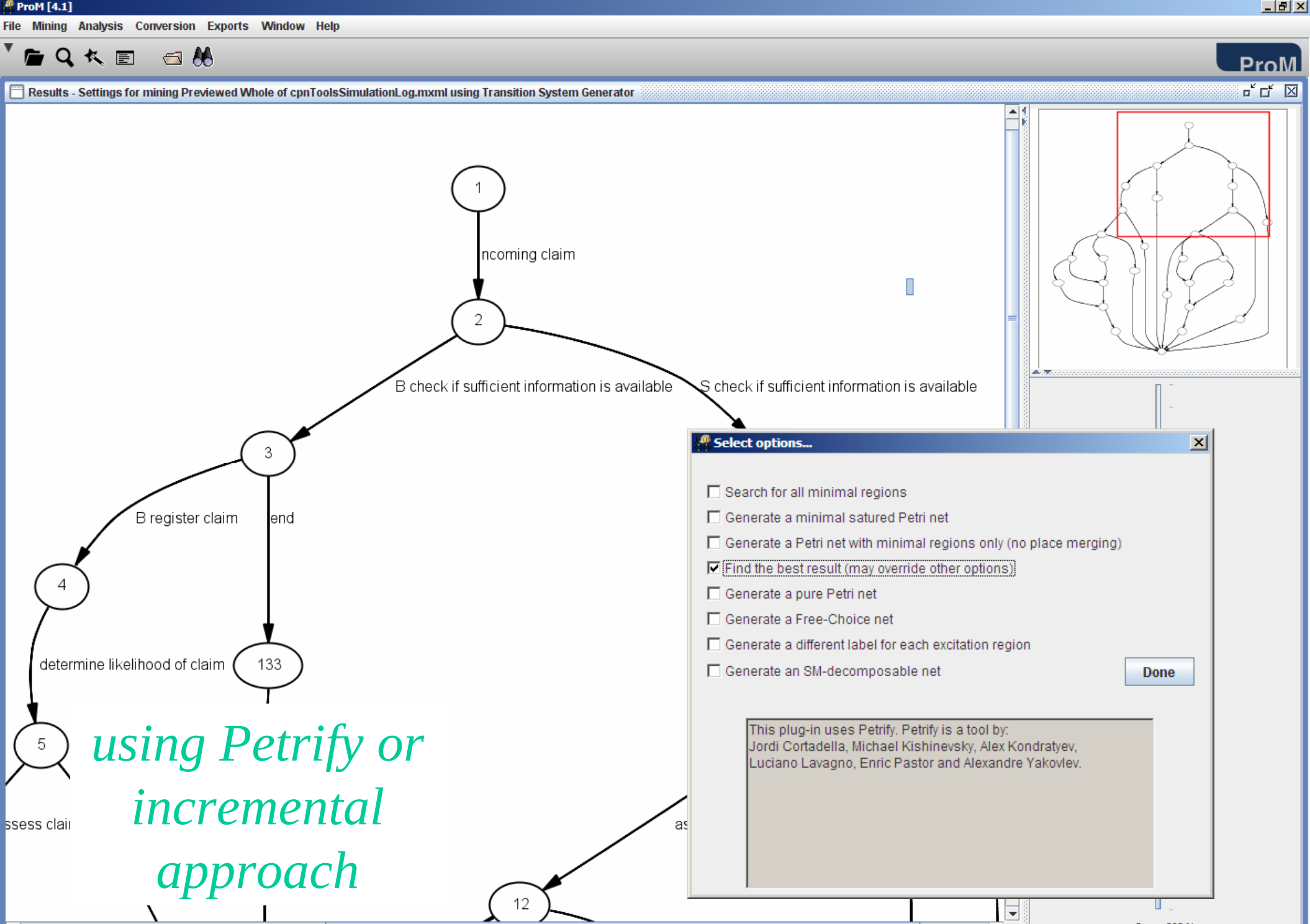


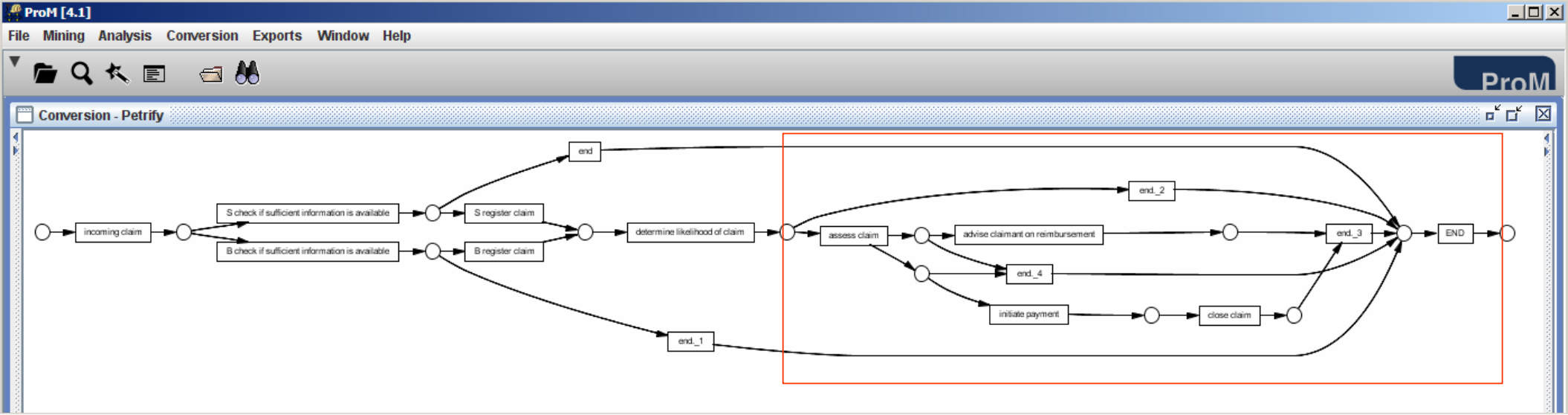


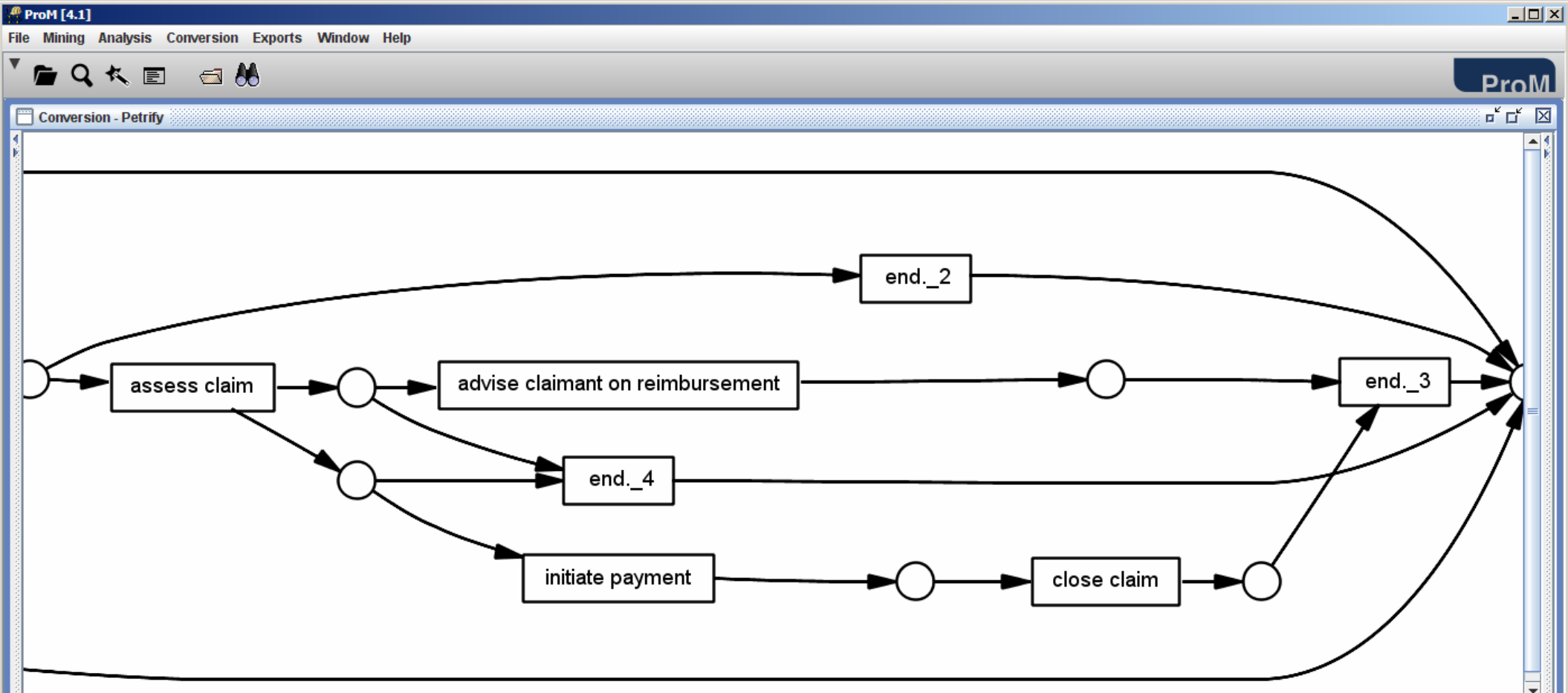
Zoom: 218 %

Region-based miner

fitting model, time-consuming, tends to overfit, …


ProM [4.1]
File Mining Analysis Conversion Exports Window Help
Results - Settings for mining Previewed Whole of cpnToolsSimulationLog.mxml using Transition System Generator
1
ncoming claim
2
B check if sufficient information is available
S check if sufficient information is available
3
B register claim
end
4
determine likelihood of claim
133
5
ssess clai
as
12
using Petrify or incremental approach
Select options...
Search for all minimal regions
Generate a minimal satured Petri net
Generate a Petri net with minimal regions only (no place merging)
Find the best result (may override other options)
Generate a pure Petri net
Generate a Free-Choice net
Generate a different label for each excitation region
Generate an SM-decomposable net
Done
This plug-in uses Petrify. Petrify is a tool by:
Jordi Cortadella, Michael Kishinevsky, Alex Kondratyev,
Luciano Lavagno, Enric Pastor and Alexandre Yakovlev.


ProM [4.1]
File Mining Analysis Conversion Exports Window Help
ProM
Conversion - Petrify
incoming claim
S check if sufficient information is available
B check if sufficient information is available
S register claim
B register claim
end
determine likelihood of claim
assess claim
advise claimant on reimbursement
end_2
end_3
end_4
initiate payment
close claim
end_1
END


ProM [4.1]
File Mining Analysis Conversion Exports Window Help
ProM
Conversion - Petrify
end._2
assess claim
advise claimant on reimbursement
end._3
end._4
initiate payment
close claim

Fuzzy miner

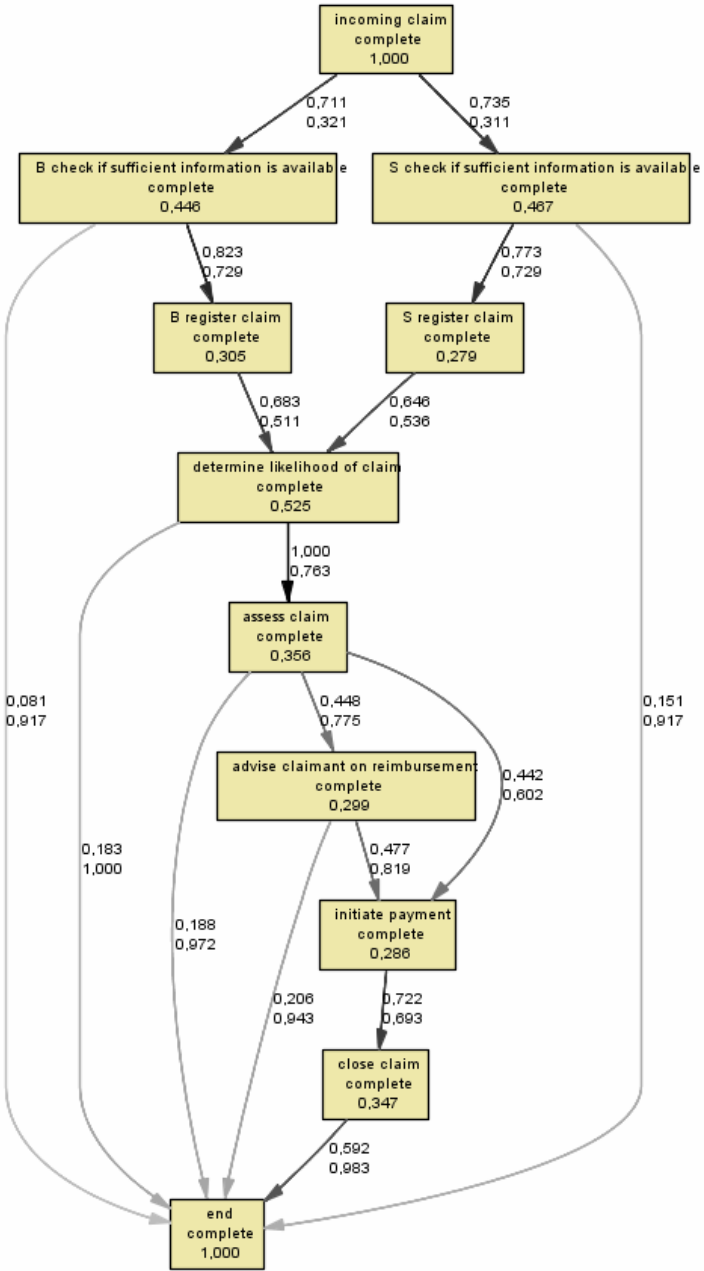
robust, dealing with Spaghetti, tunable, …

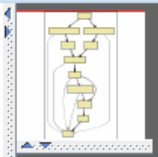
ProM [4.1]

File Mining Analysis Conversion Exports Window Help

Results - Settings for mining Filtered cpnToolsSimulationLog.mxml using Fuzzy Miner

Graph | Unary metrics | Binary metrics





Concurrency filter

Edge filter

Edge transformer

Best edges

Fuzzy edges

S/C ratio

Cutoff

0.750

1,000

ignore self-loops

interpret absolute

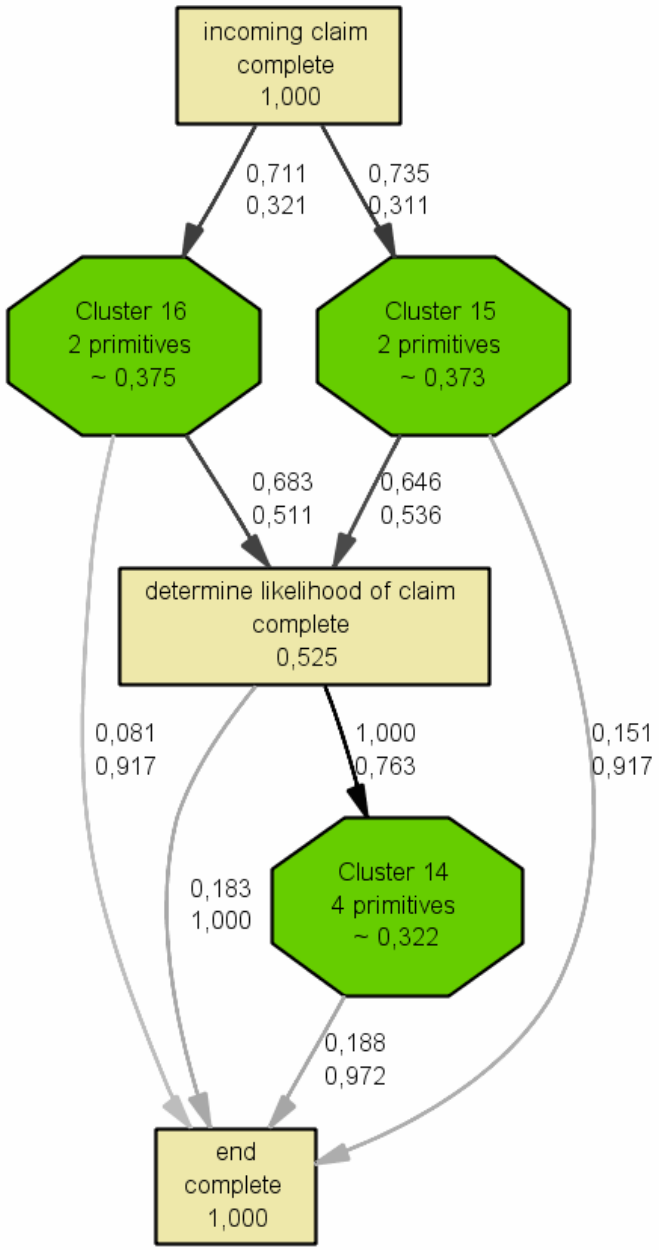
Node filter

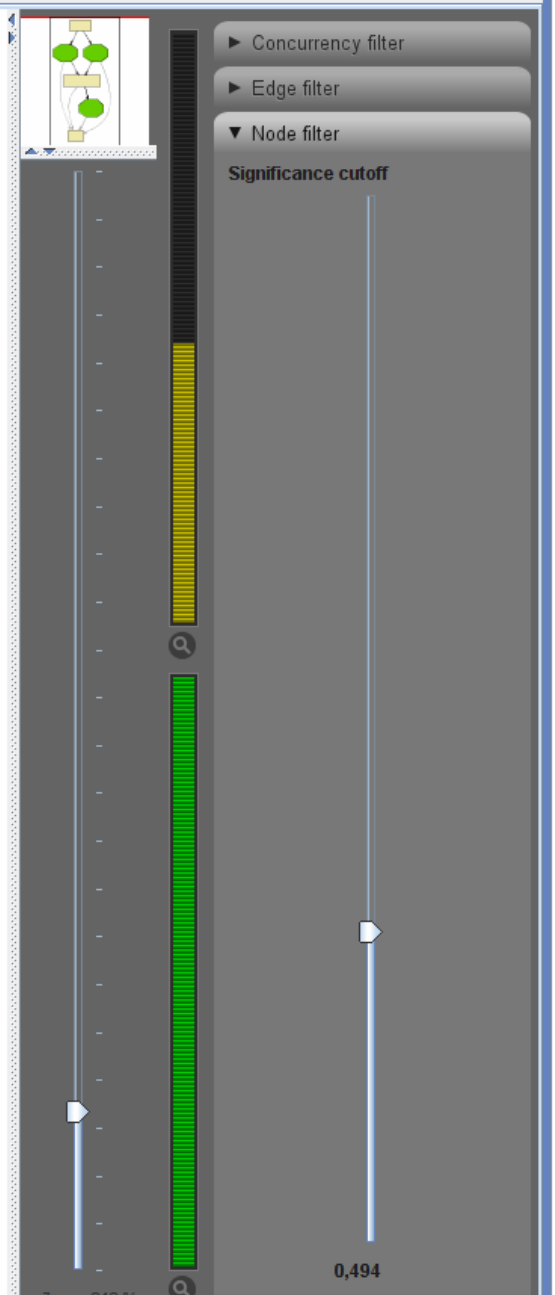
ProM [4.1]

File Mining Analysis Conversion Exports Window Help

Results - Settings for mining Filtered cpnToolsSimulationLog.mxml using Fuzzy Miner

Graph | Unary metrics | Binary metrics



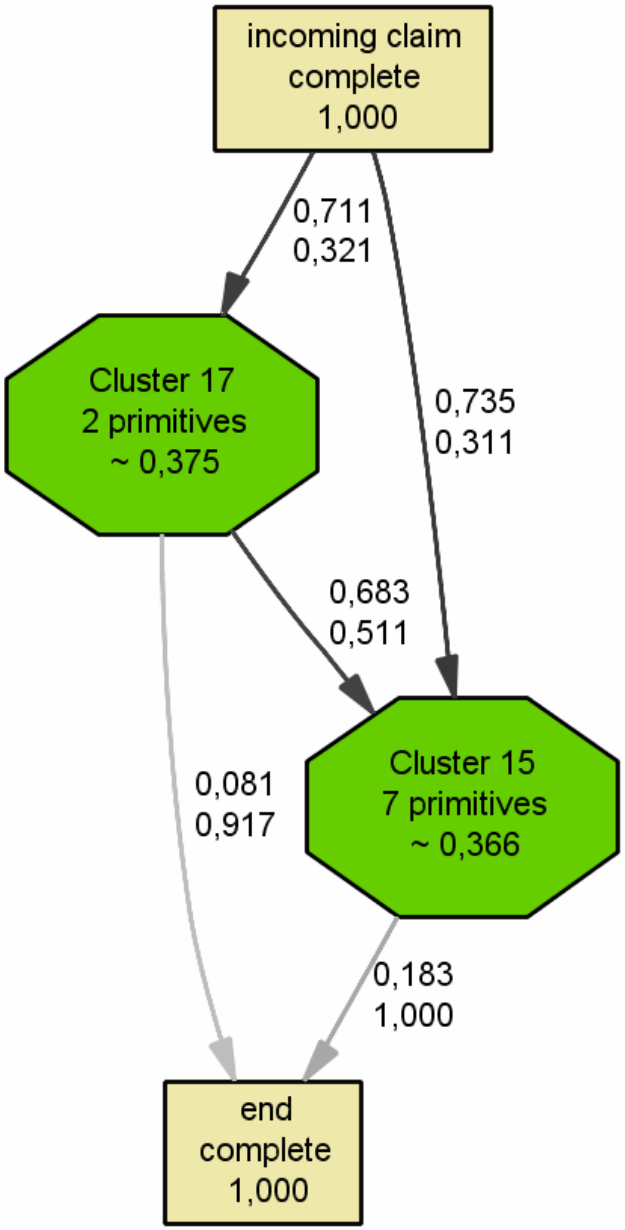


ProM [4.1]

File Mining Analysis Conversion Exports Window Help

Results - Settings for mining Filtered cpnToolsSimulationLog.mxml using Fuzzy Miner

Graph | Unary metrics | Binary metrics



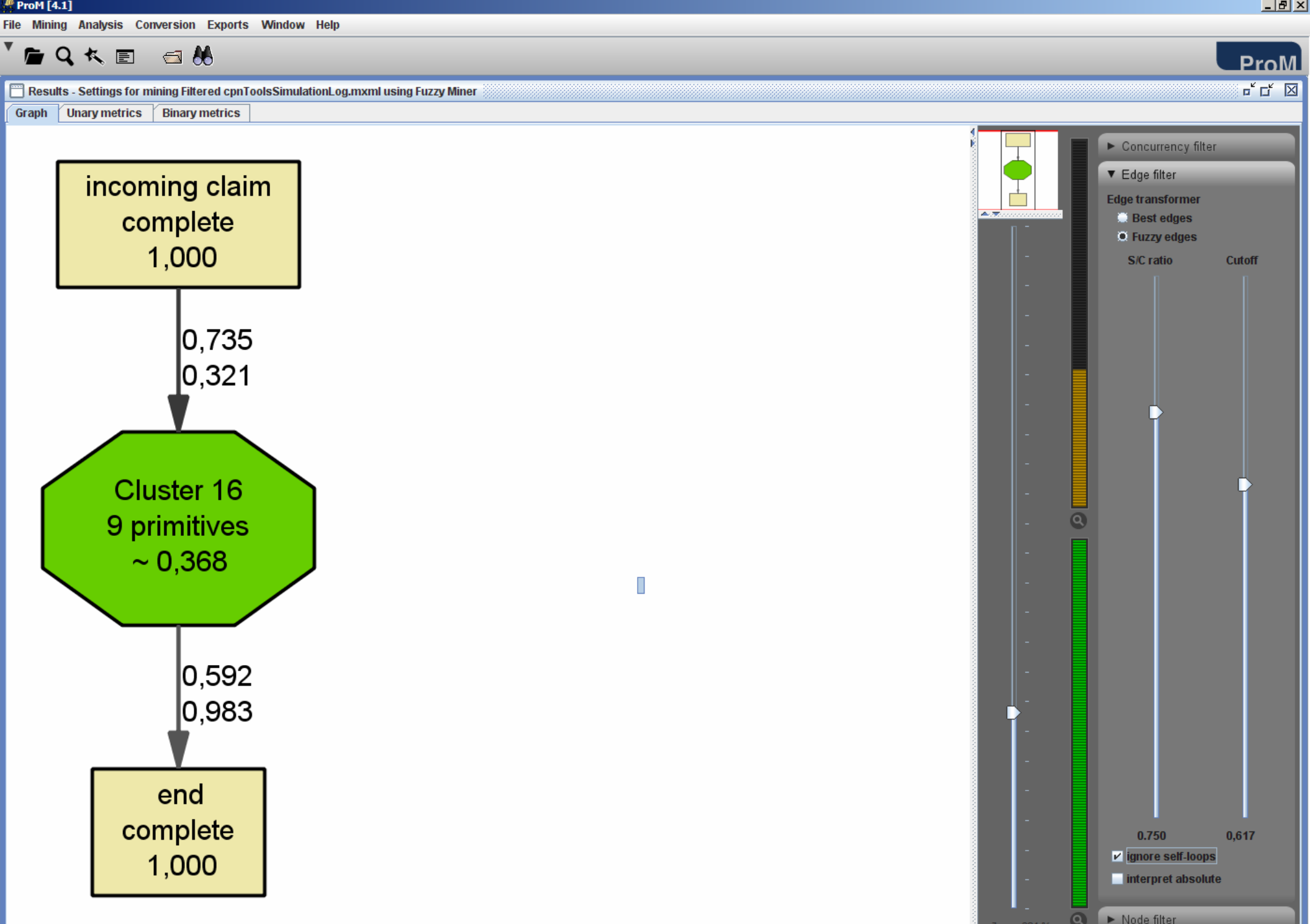
Concurrency filter

Edge filter

Node filter

Significance cutoff

0,797


ProM [4.1]
File Mining Analysis Conversion Exports Window Help
ProM
Results - Settings for mining Filtered cpnToolsSimulationLog.mxml using Fuzzy Miner
Graph
Unary metrics
Binary metrics
incoming claim
complete
1,000
0,735
0,321
Cluster 16
9 primitives
~ 0,368
0,592
0,983
end
complete
1,000
Concurrency filter
Edge filter
Edge transformer
Best edges
Fuzzy edges
S/C ratio
Cutoff
0.750
0,617
ignore self-loops
interpret absolute
Node filter

Social network miner

using references to people, groups, machines, …

ProM [4.1]

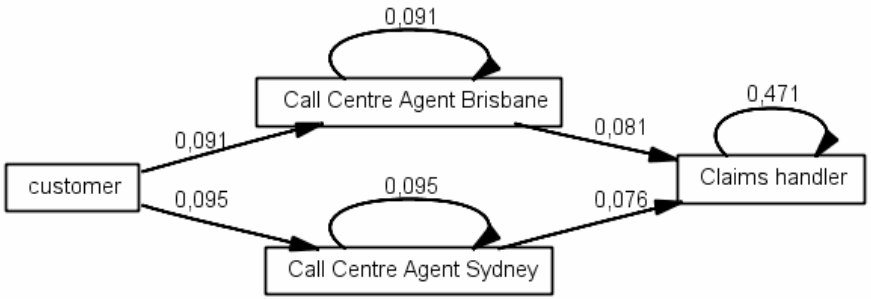
File Mining Analysis Conversion Exports Window Help

Results - Settings for mining Filtered cpnToolsSimulationLog.mxml using Social network miner

	Call Centre ...	Call Centre ...	Claims han...	customer
Call Centre...	0.0911741...	0.0	0.0808375...	0.0
Call Centre...	0.0	0.0949907...	0.0755367...	0.0
Claims ha...	0.0	0.0	0.4712960...	0.0
customer	0.0911741...	0.0949907...	0.0	0.0

threshold : 0

Remove isolated nodes



Zoom

Analysis - Analyze Social Network

Show role nodes
Show org unit nodes
Vertex size
Size property: Frequency
Vertex degree ratio stretch
Edge weight

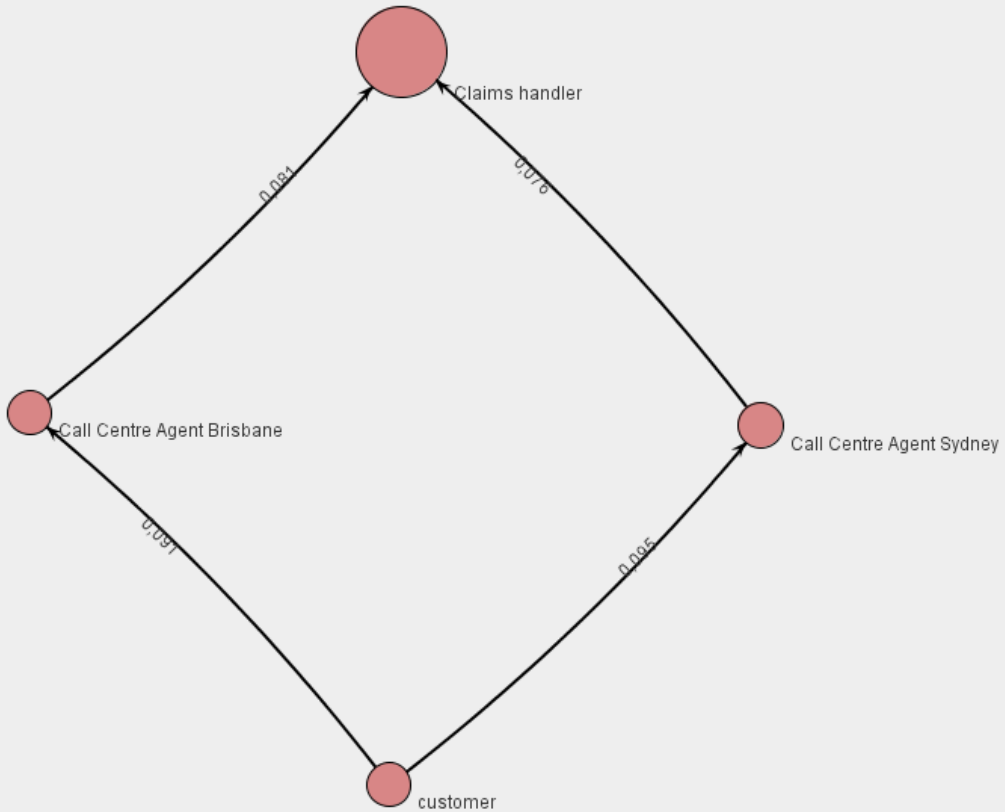
Layout: KKLayout

Mouse Mode: Transforming

Group Clusters

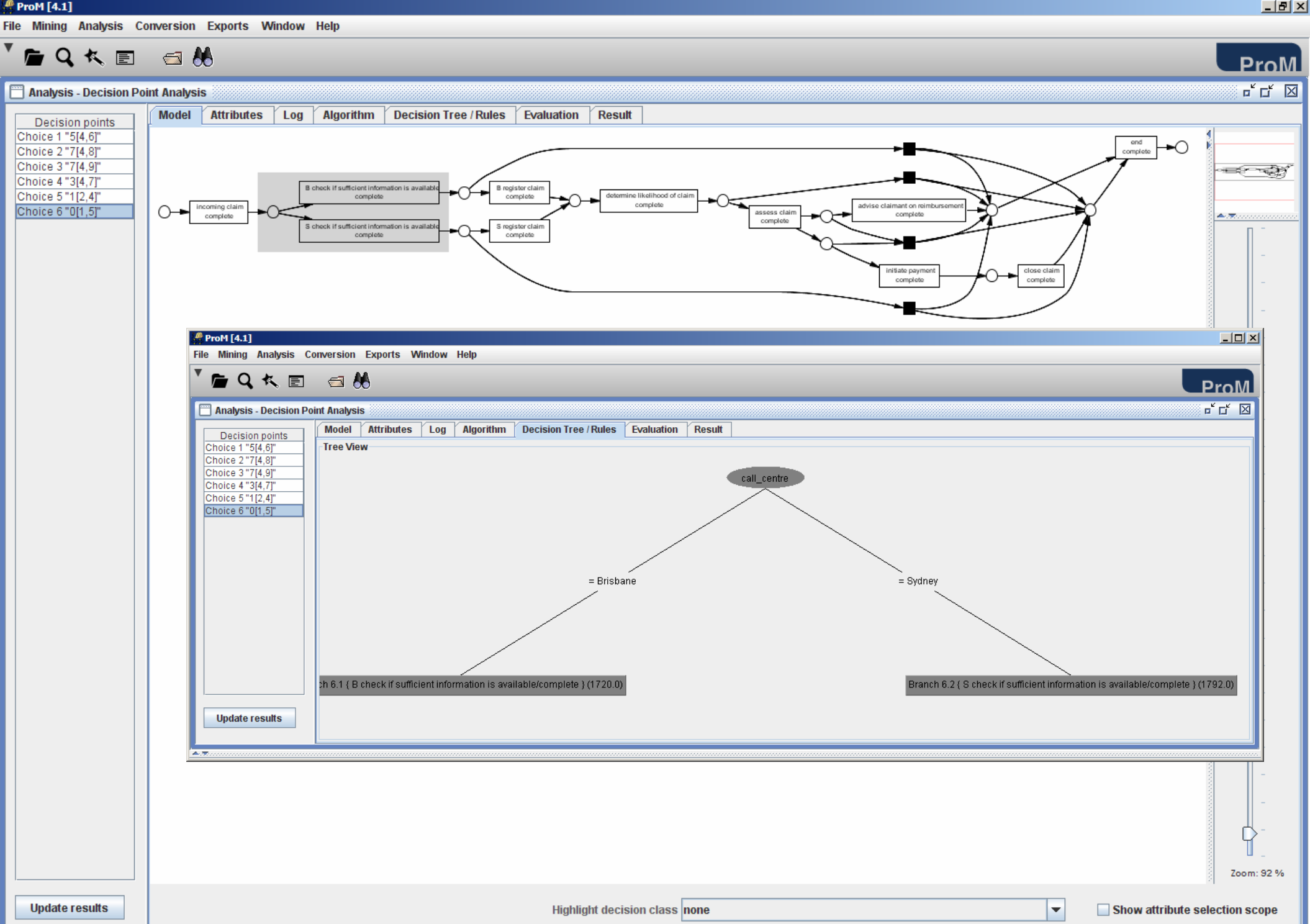
Weight | Betweenness

Edges removed for clusters: 0



Decision miner

using data in the log, …


ProM [4.1]
File Mining Analysis Conversion Exports Window Help
Analysis - Decision Point Analysis
Decision points
Choice 1 "5[4,6]"
Choice 2 "7[4,8]"
Choice 3 "7[4,9]"
Choice 4 "3[4,7]"
Choice 5 "1[2,4]"
Choice 6 "0[1,5]"
Model Attributes Log Algorithm Decision Tree / Rules Evaluation Result
incoming claim complete
B check if sufficient information is available complete
S check if sufficient information is available complete
B register claim complete
S register claim complete
determine likelihood of claim complete
assess claim complete
advise claimant on reimbursement complete
initiate payment complete
close claim complete
end complete
ProM [4.1]
File Mining Analysis Conversion Exports Window Help
Analysis - Decision Point Analysis
Decision points
Choice 1 "5[4,6]"
Choice 2 "7[4,8]"
Choice 3 "7[4,9]"
Choice 4 "3[4,7]"
Choice 5 "1[2,4]"
Choice 6 "0[1,5]"
Update results
Model Attributes Log Algorithm Decision Tree / Rules Evaluation Result
Tree View
call_centre
= Brisbane
= Sydney
ch 6.1 { B check if sufficient information is available/complete } (1720.0)
Branch 6.2 { S check if sufficient information is available/complete } (1792.0)
Update results
Zoom: 92 %
Highlight decision class none
Show attribute selection scope

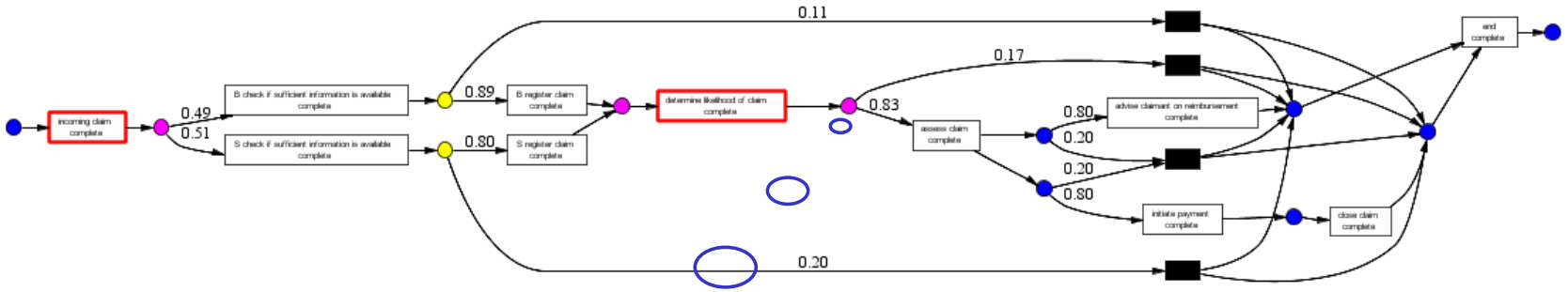
Performance miner

using timestamps in log, …

ProM [4.1]

File Mining Analysis Conversion Exports Window Help

Analysis - Performance Analysis with Petri net



bottle-necks

throughput time

flow time from A to B

Process information:

Total number selected:

3512 cases

Number fitting:

3512 cases

Arrival rate:

0,12 cases per second

	Throughput time (seconds)
avg	11115,54
min	0,0
max	40704,0
stdev	8906,98
fast 25...	1379,19
slow 2...	23817,24
norma...	9632,87

Change Percentages | Export Time-Metrics

Performance information of the selected transitions:

Frequency: 2950 cases

	Time in between (seconds)
avg	12248,87
min	53,0
max	39706,0
stdev	8381,14

Waiting time:
High
Medium
Low
Settings

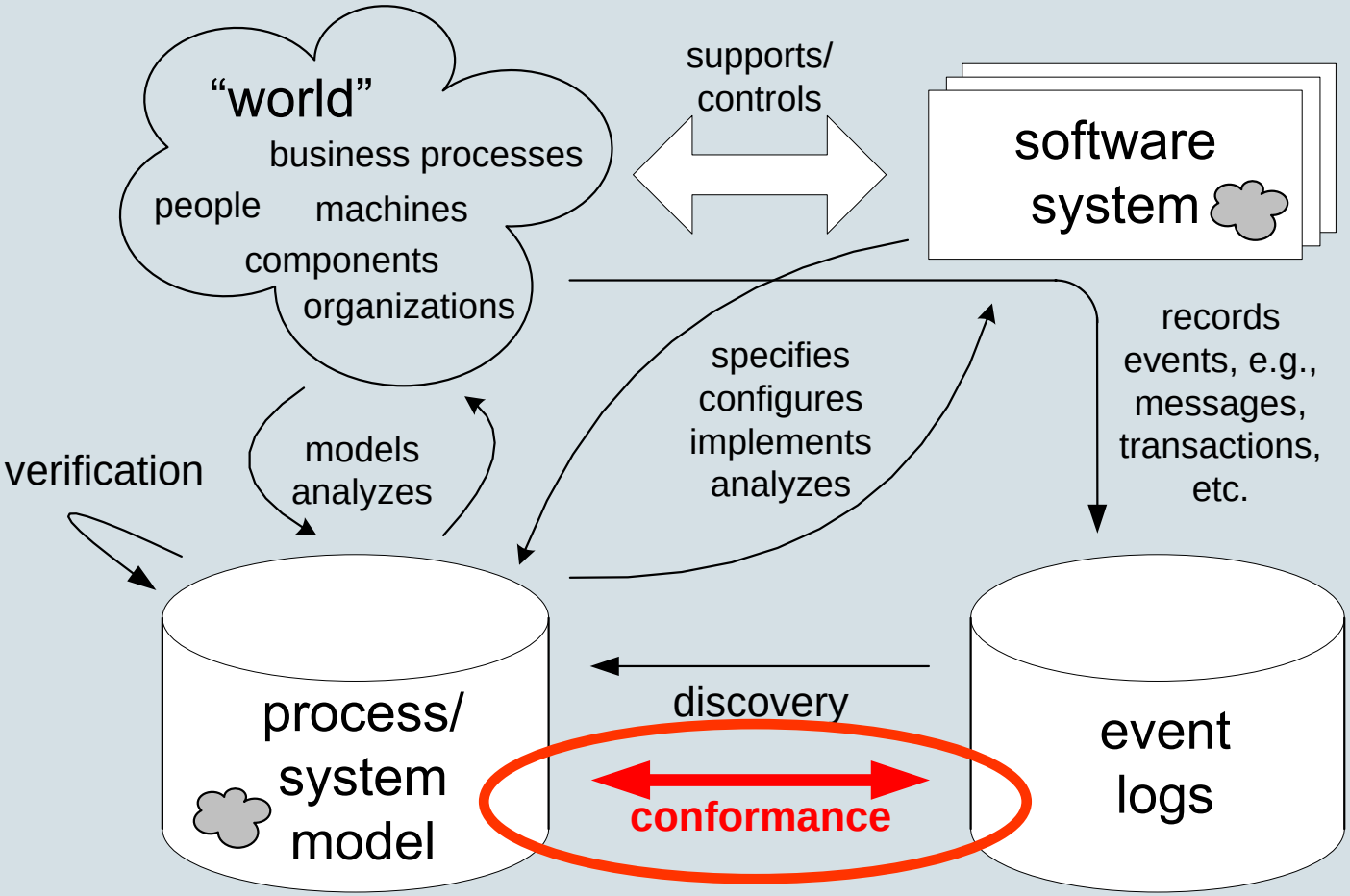
Selected:

Transition - incoming claim c...

and:

Transition - determine likeliho...

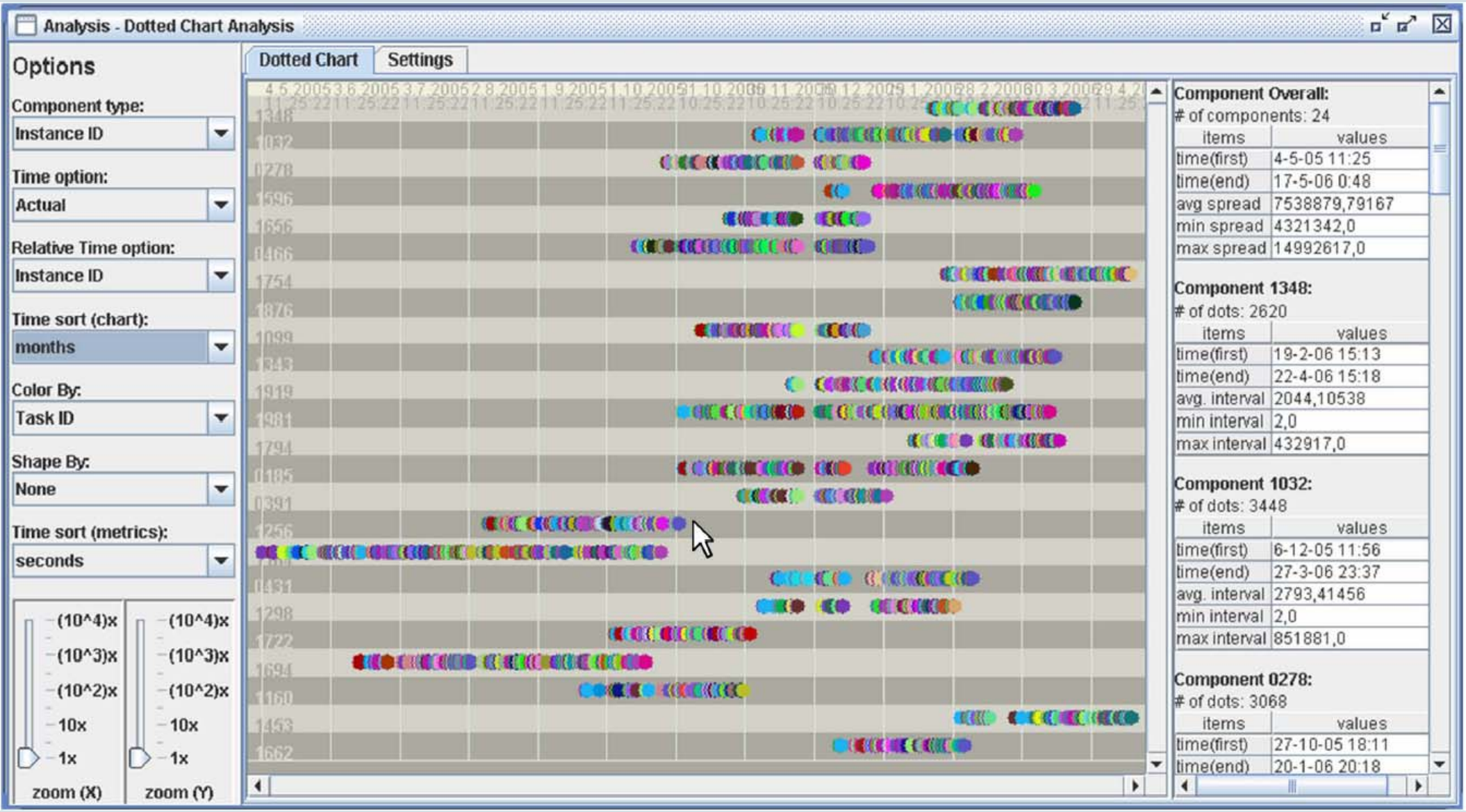
Conformance Checking



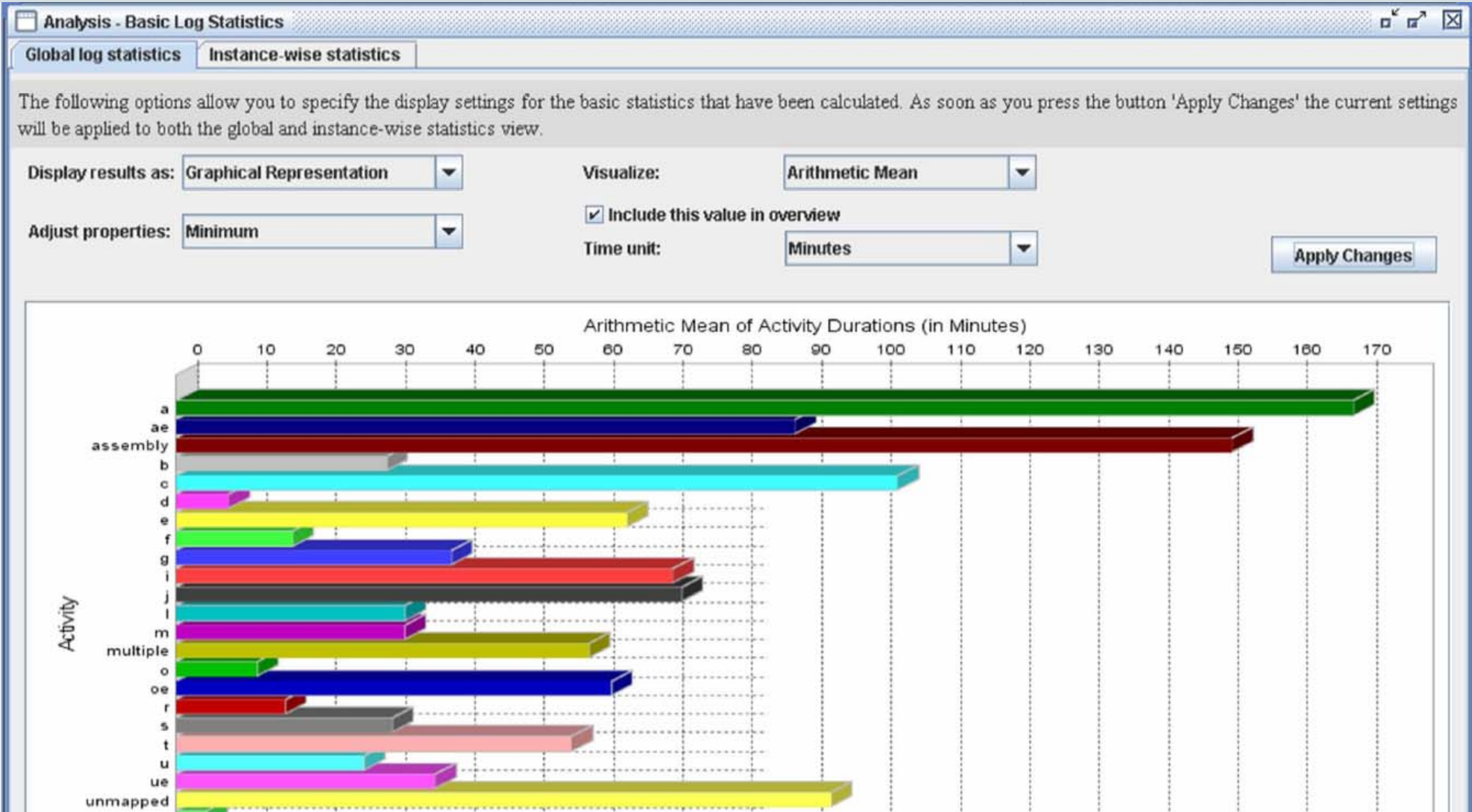
Challenging example: Test data from high-tech system manufacturer (ASML)

- 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)

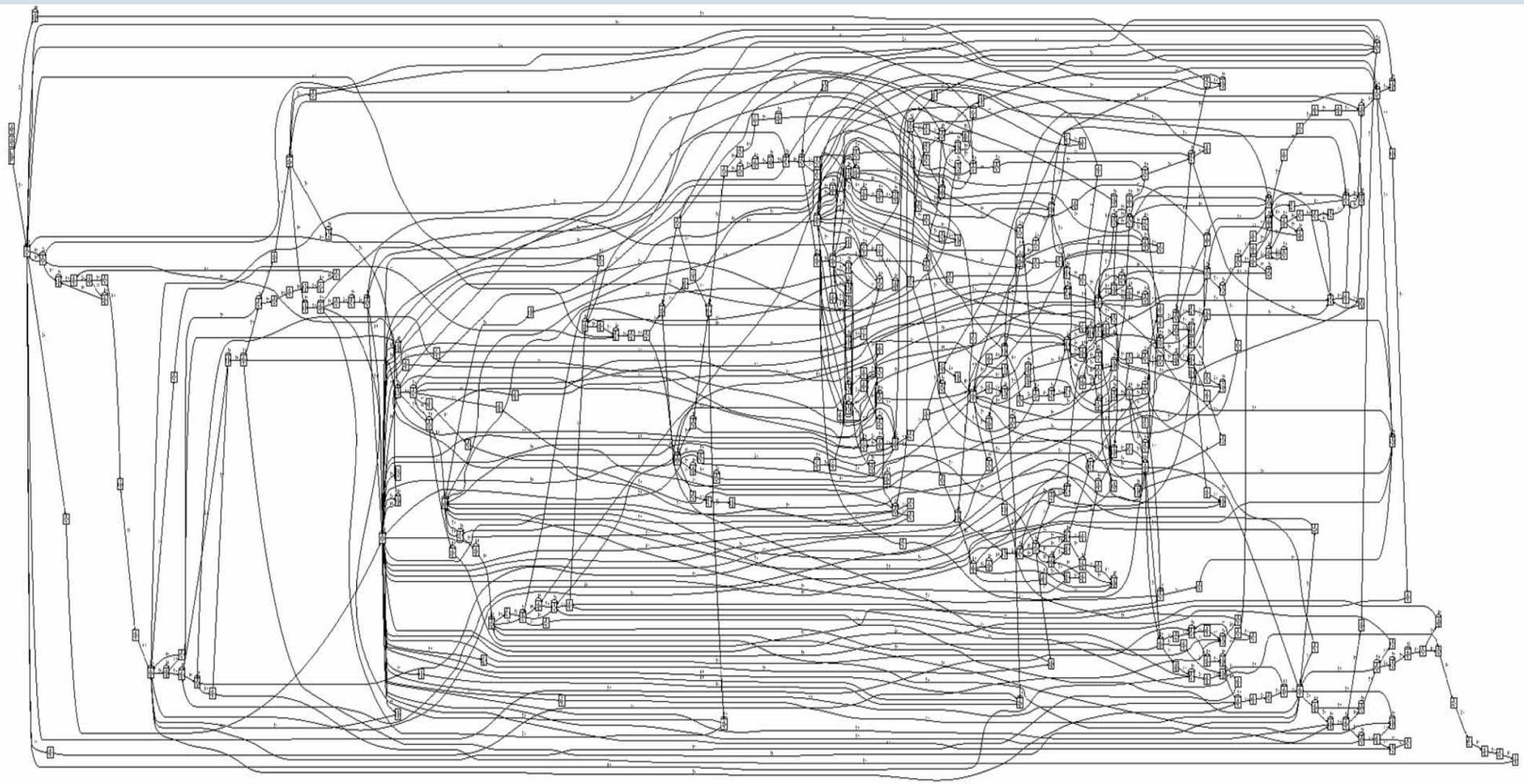
Helicopter view



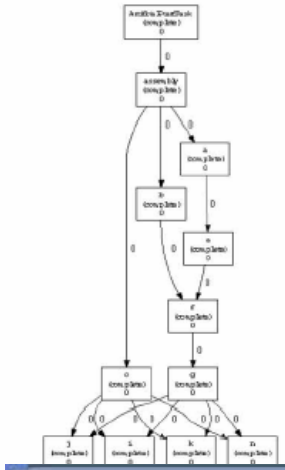
Average time spent in job-steps (aggregated events)

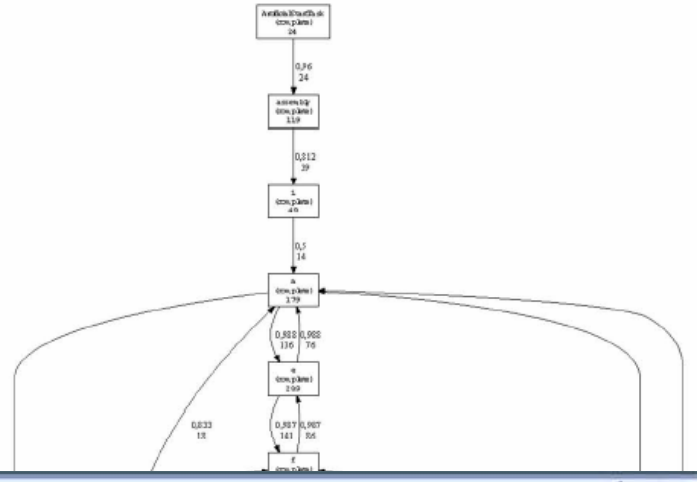


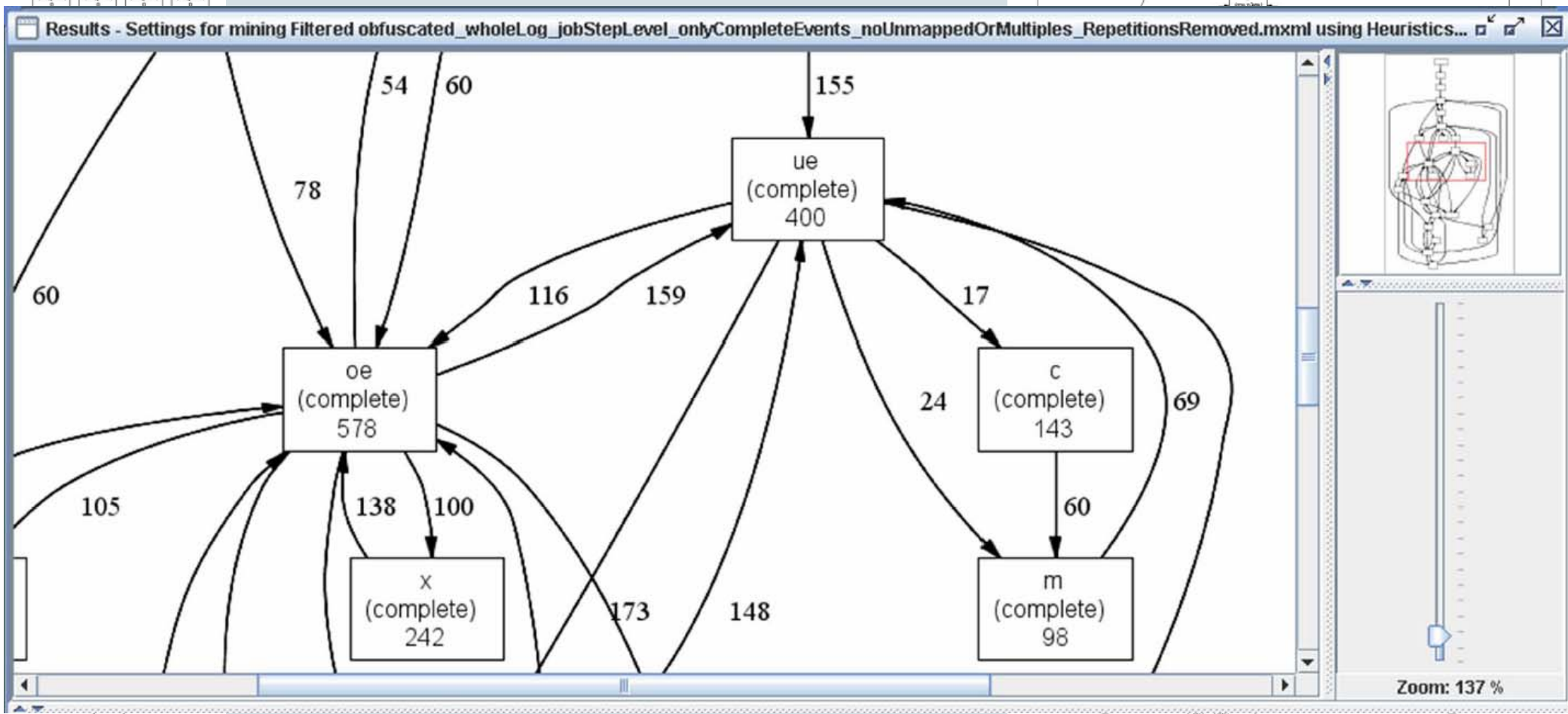
Mining just the complete events (# 360)…



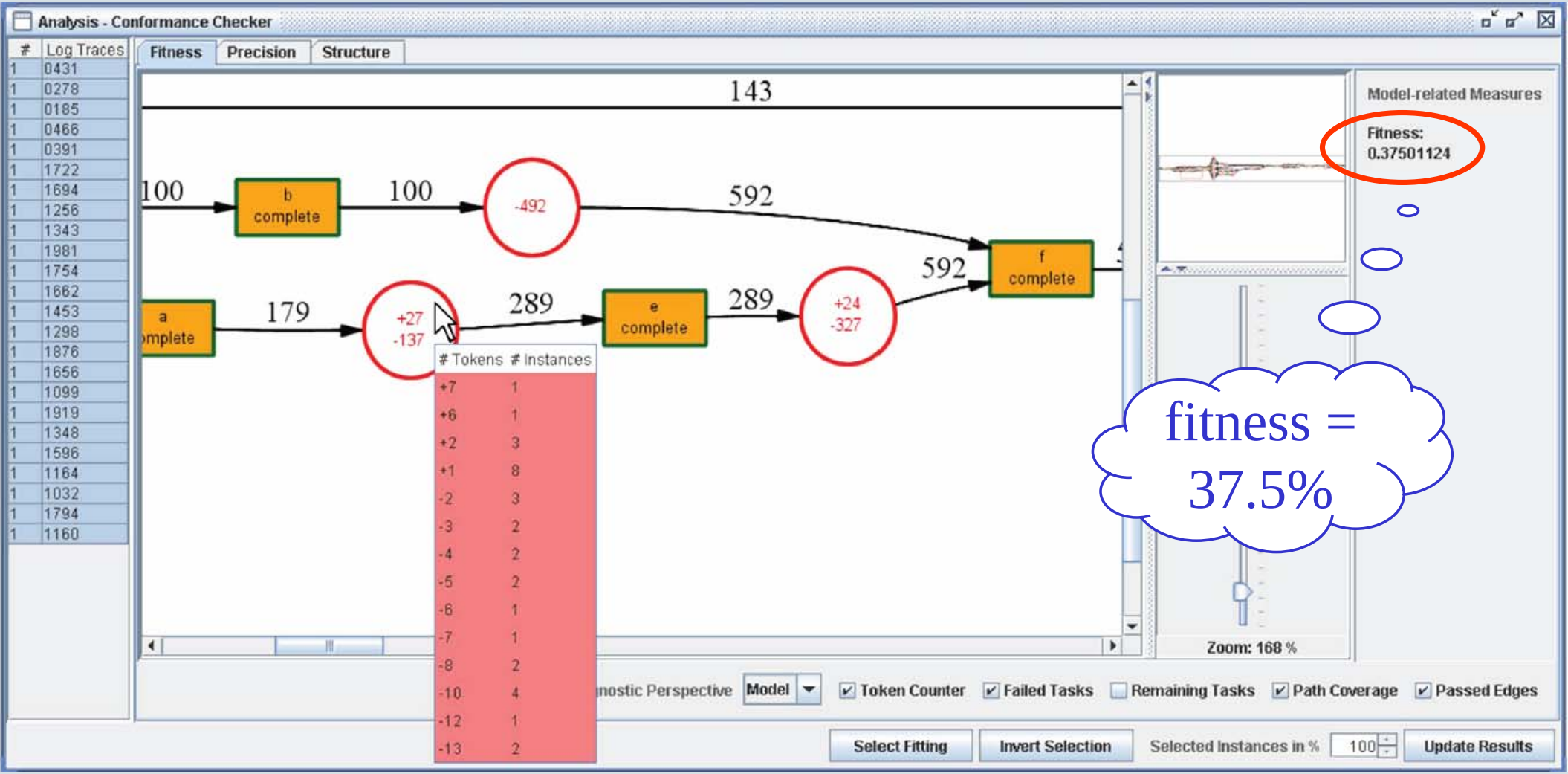
Job step level





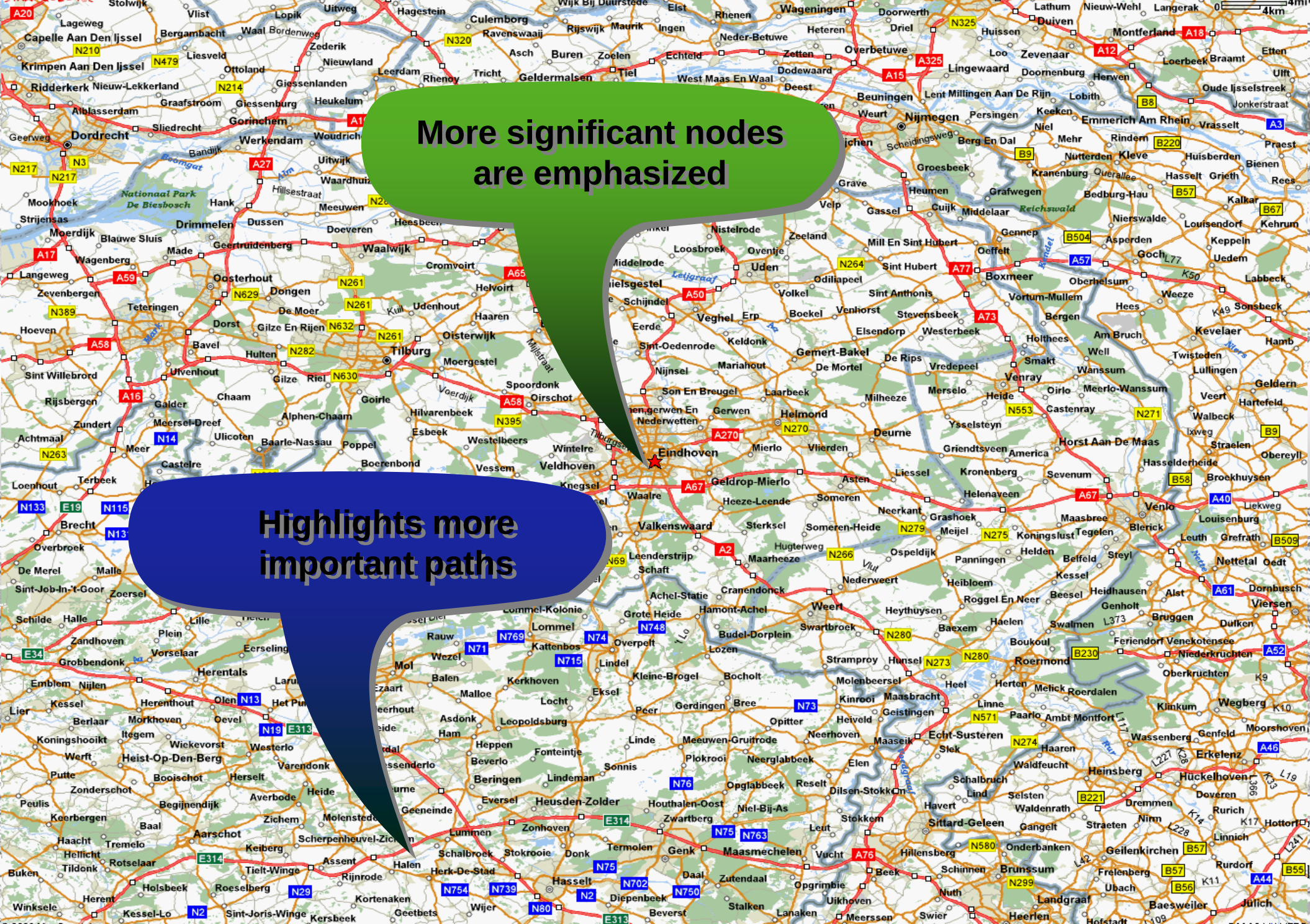


Conformance checker (reference model – job steps)



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

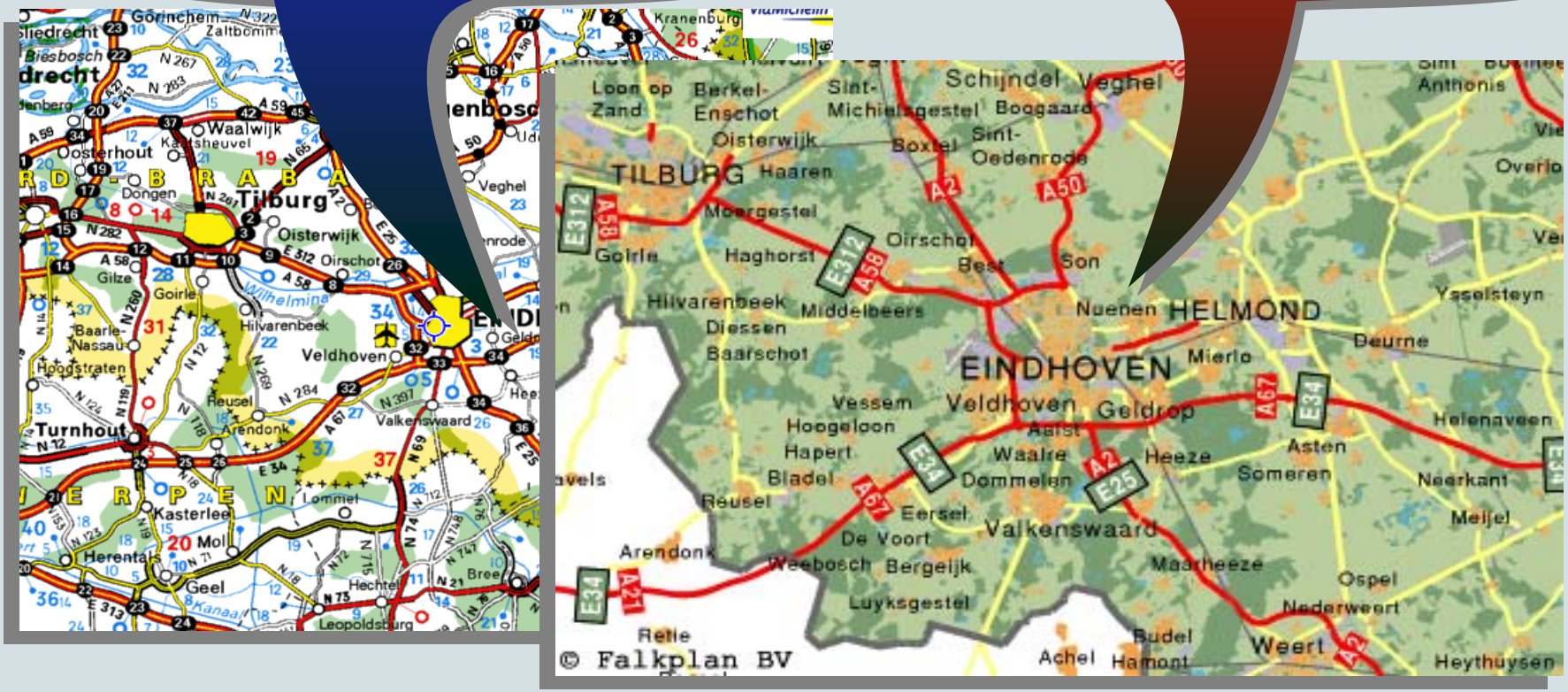
More to learn from maps...

Aggregation

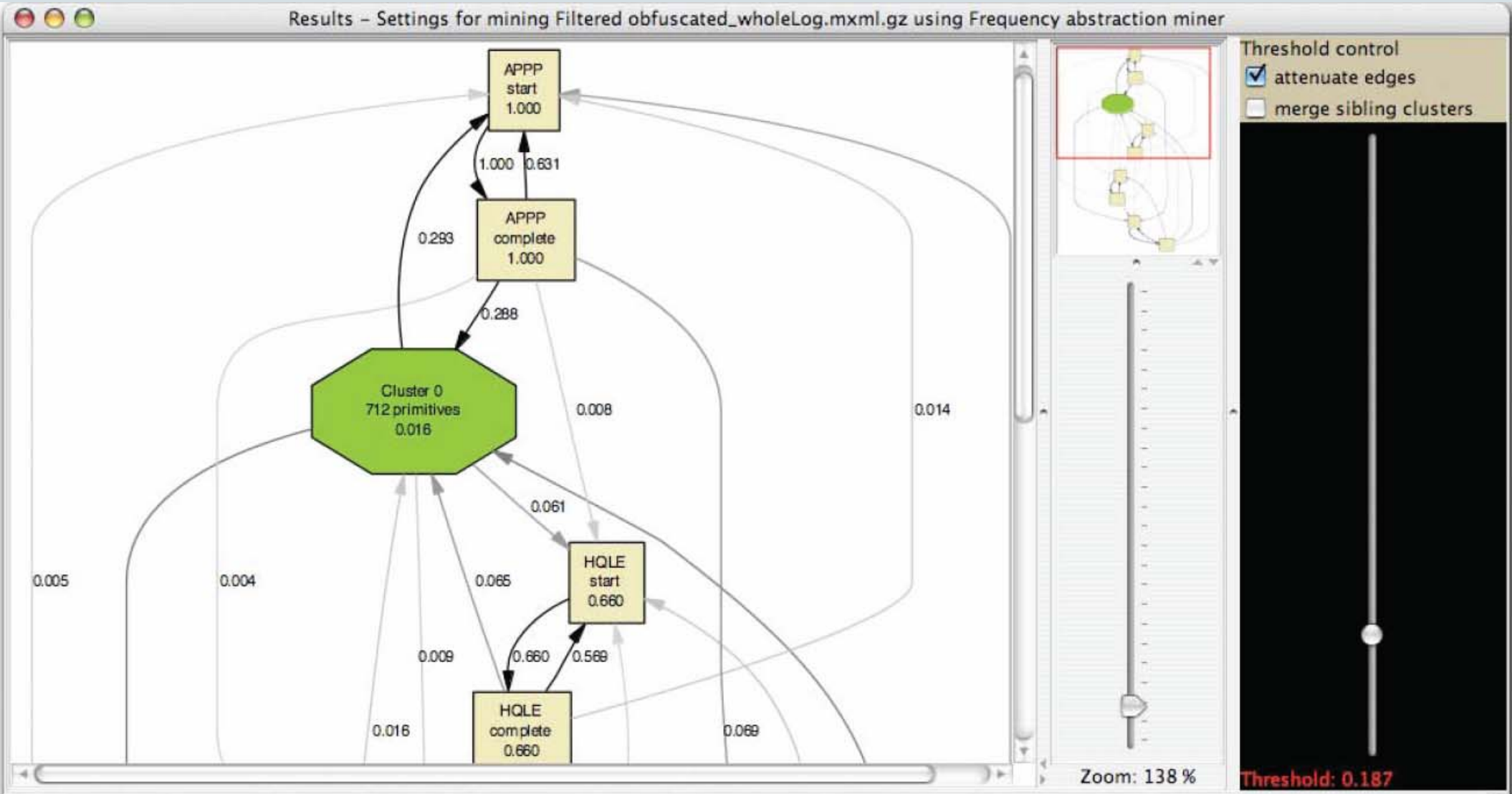
Clustering of coherent, less significant structures

Abstraction

Removing isolated, less significant structures



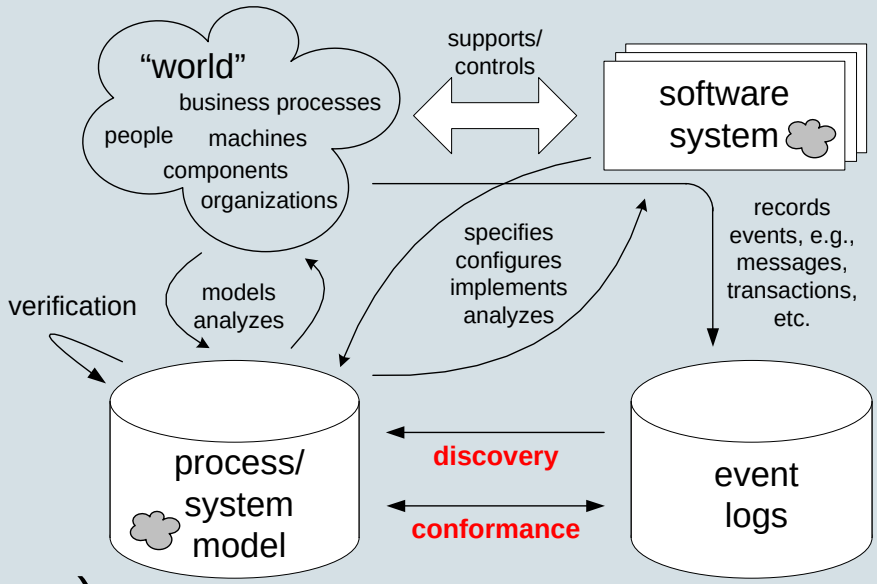
ProM's Frequency abstraction miner applied to ASML log

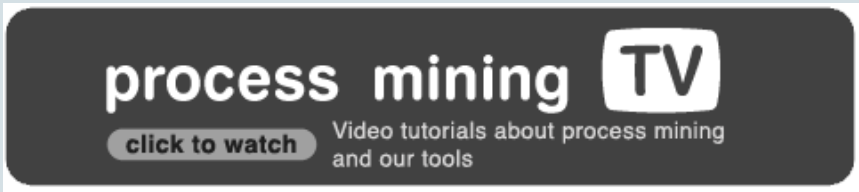


ProM analyze your MXML logs

Conclusion

- The existence of event data enables a wide variety of **process mining** techniques: discovery and conformance.
- A **reality check** for people that are involved in process modeling.
- Interesting **challenges** for Petri net researchers.
- **ProM** supports this (+185 plug-ins).
- **ProM** also supports:
  - model conversions (EPC, YAWL, BPEL, …)
  - verification
  - automatic generation of CPN models



Relevant WWW sites

- **http://www.processmining.org**
- http://promimport.sourceforge.net
- http://prom.sourceforge.net
- http://www.workflowpatterns.com
- http://www.workflowcourse.com
- http://www.win.tue.nl/is/
- http://is.tm.tue.nl/staff/wvdaalst