Process Mining: Organizational and Conformance Mining Algorithms

Ana Karla Alves de Medeiros

Eindhoven University of Technology
Department of Information Systems
a.k.medeiros@tue.nl

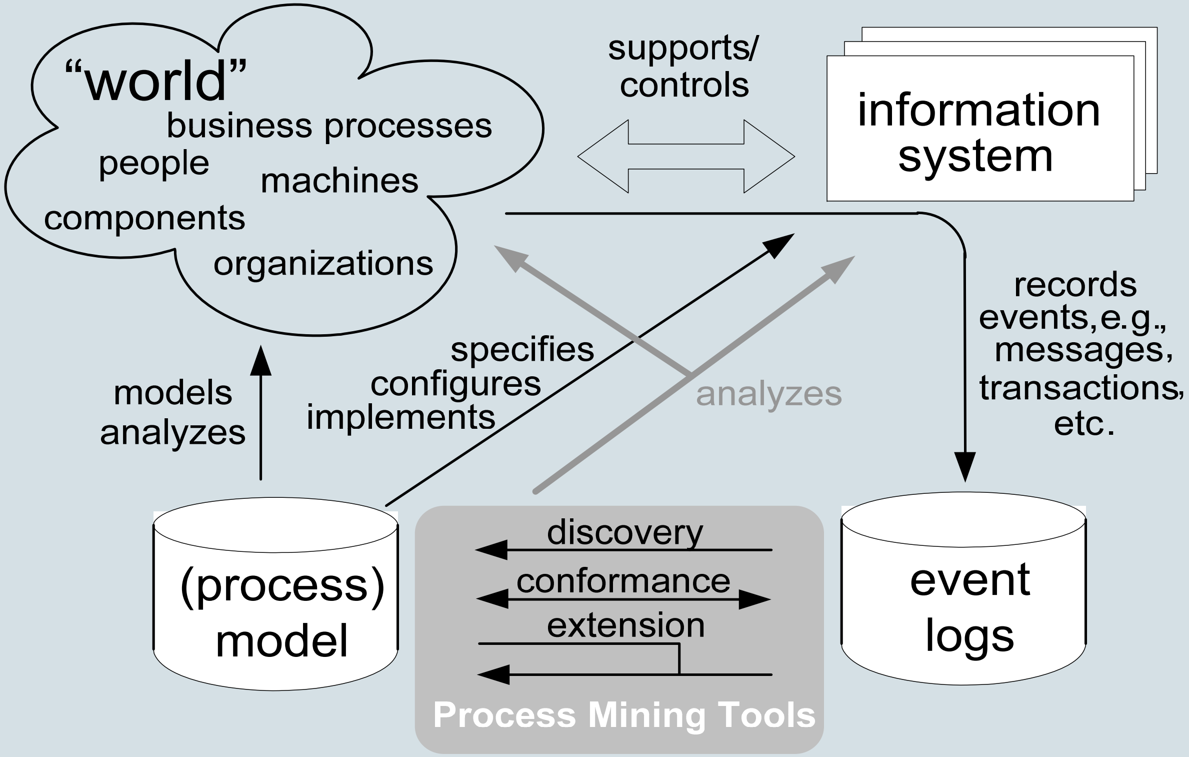
Process Mining

- Short Recap
- Discovery Techniques (Part 2)
  - Organizational Model
  - Social Network
- Conformance Techniques
  - Conformance Checker
  - LTL- Checker
- Summary
- Announcements

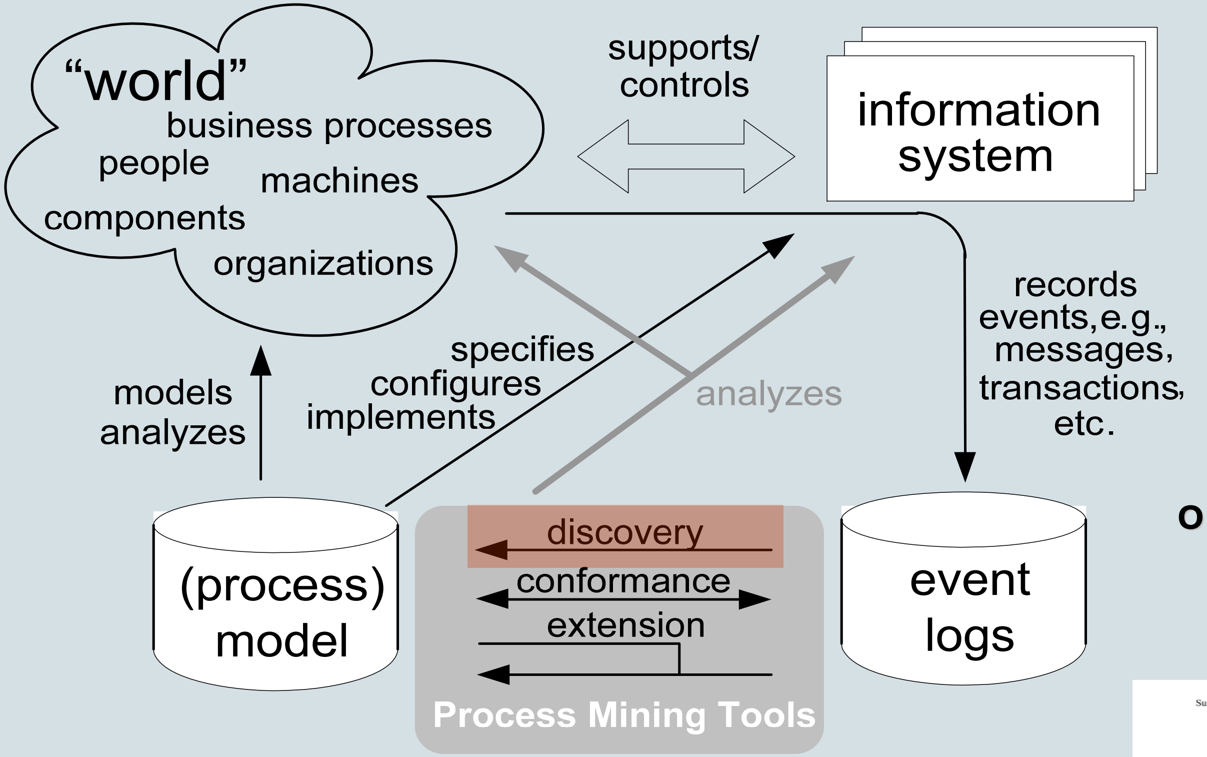
Process Mining

- Short Recap
- Discovery Techniques (Part 2)
  - Organizational Model
  - Social Network
- Conformance Techniques
  - Conformance Checker
  - LTL- Checker
- Summary
- Announcements

Types of Algorithms



Types of Algorithms



"world"
business processes
people
machines
components
organizations

supports/
controls

information
system

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

models
analyzes

specifies
configures
implements

analyzes

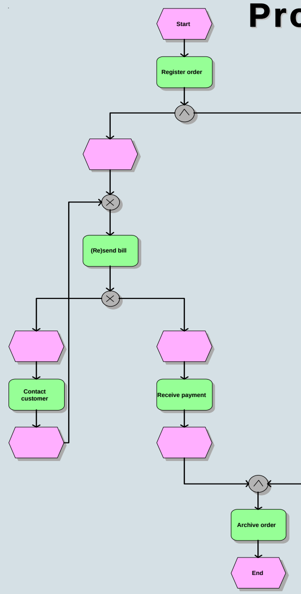
(process)
model

discovery
conformance
extension

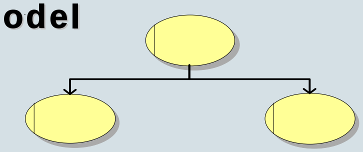
Process Mining Tools

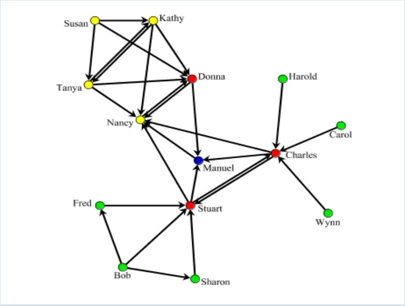
event
logs

Process Model



Organizational Model

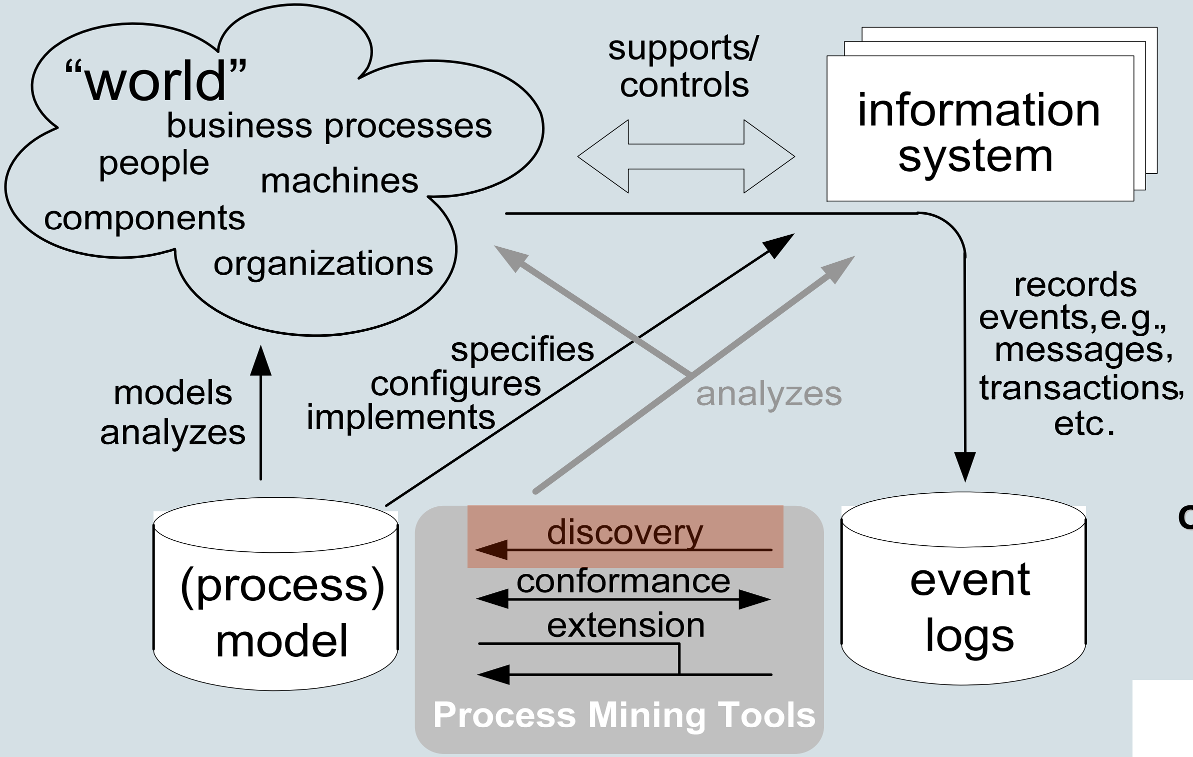


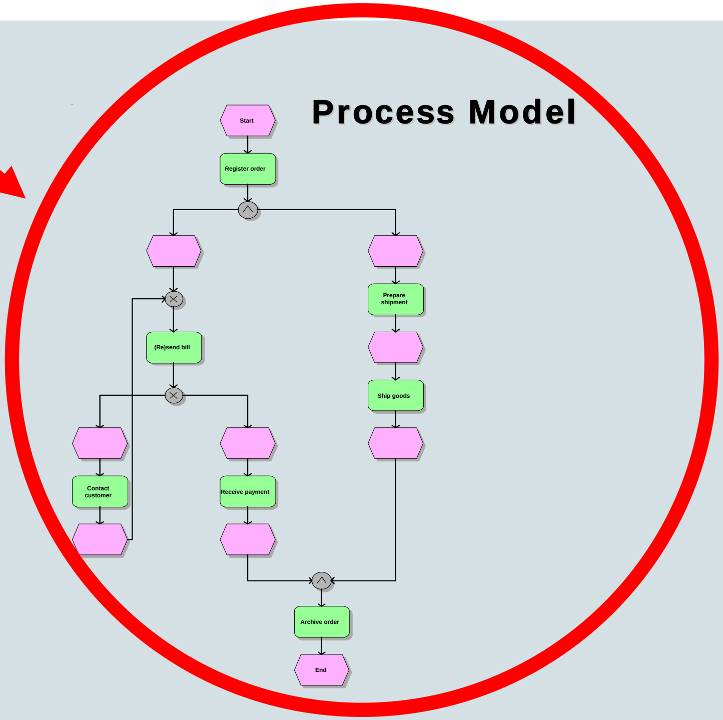


Social Network

Control-flow mining techniques

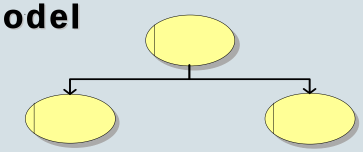
Types of Algorithms

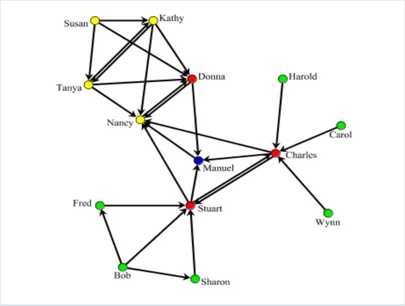




Process Model

Organizational Model





Social Network

Main Points Lecture 3

- The notion of a process instance is crucial!
- Ordering of tasks is the basic information
- Frequencies are important to handle noise
- Local approaches
  - α-algorithm, Heuristics Miner
- Global approaches
  - Genetic Miner and Fuzzy Miner

Do you still remember why?

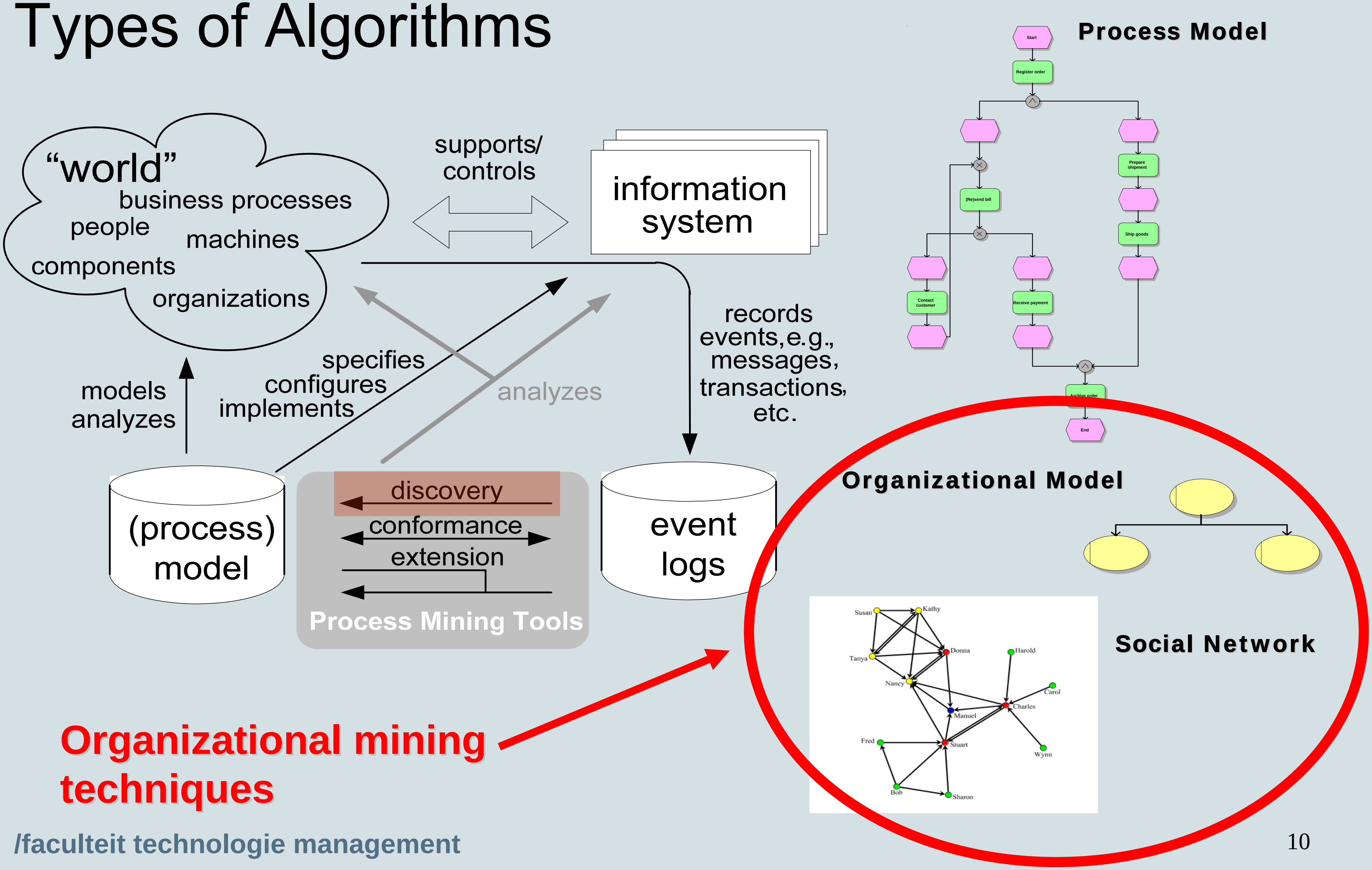
Process Mining

- Short Recap
- Discovery Techniques (Part 2)
  - Organizational Model
  - Social Network
- Conformance Techniques
  - Conformance Checker
  - LTL- Checker
- Summary
- Announcements

Process Mining

- Short Recap
- Discovery Techniques (Part 2)
  - Organizational Model
  - Social Network
- Conformance Techniques
  - Conformance Checker
  - LTL- Checker
- Summary
- Announcements

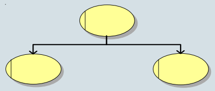
Types of Algorithms

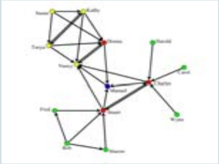


Organizational mining techniques

Organizational Mining Algorithms

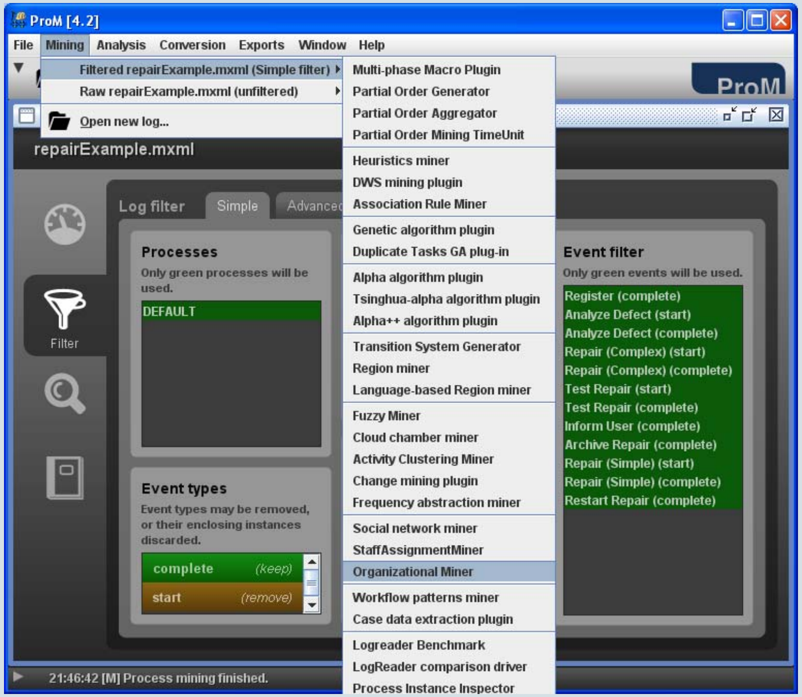
- Aid in understanding and improving social and organizational structures
- Two types of algorithms
  - Organizational Model
    - Mining of roles and teams in organizations
    - Plug-in: *Organizational Miner*
  - Social Networks
    - Discovery of relationships among originators
    - Plug-ins: *Social Network Miner* and *Analyze Social Network*





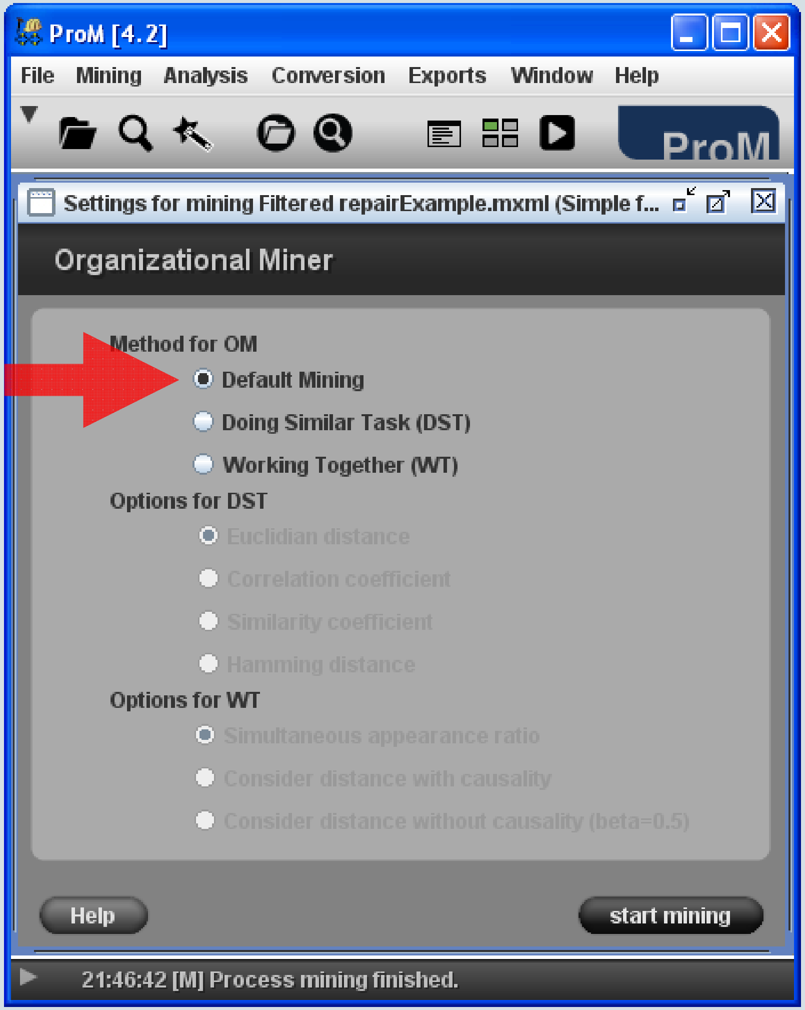
Organizational Miner

- Main idea: Which originators are executing which tasks
- Methods to mine ***roles***
  - Default mining
  - Doing Similar Tasks
- Methods to mine ***teams***
  - Working together



Organizational Miner

- Main idea: Which performers are executing which tasks
- Methods to mine ***roles***
  - ***Default mining***
  - Doing Similar Tasks
- Methods to mine ***teams***
  - Working together



Settings for mining Filtered repairExample.mxml (Simple f...

Organizational Miner

Method for OM

● Default Mining

○ Doing Similar Task (DST)

○ Working Together (WT)

Options for DST

● Euclidian distance

○ Correlation coefficient

○ Similarity coefficient

○ Hamming distance

Options for WT

● Simultaneous appearance ratio

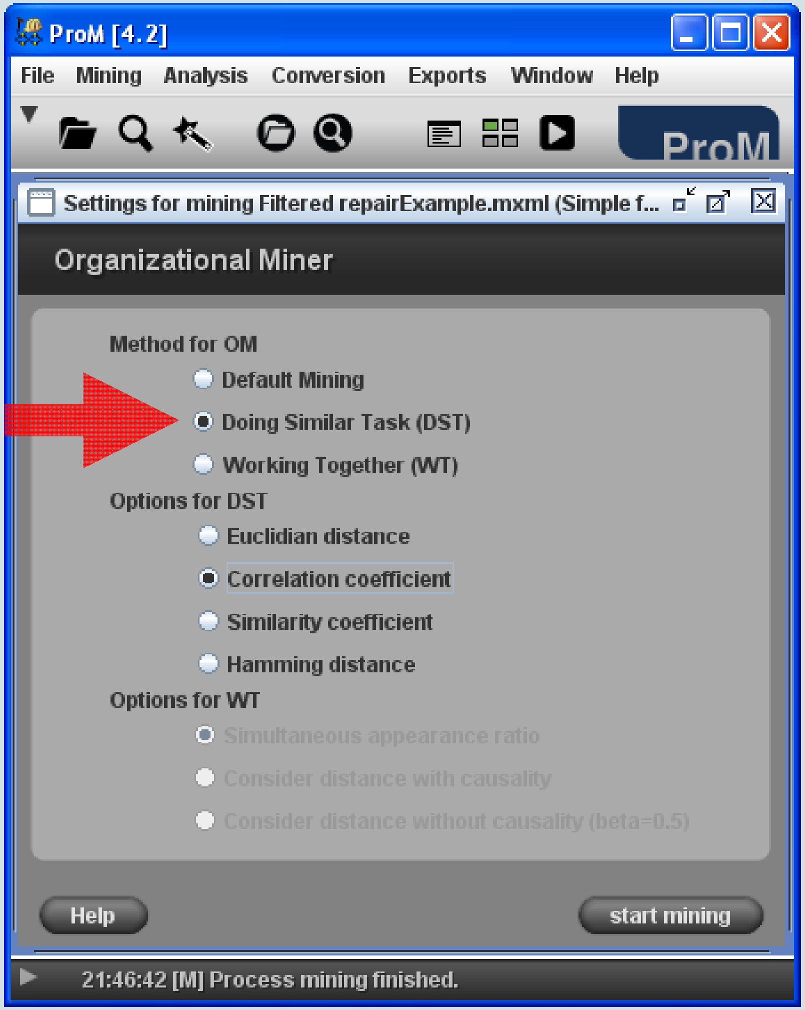
○ Consider distance with causality

○ Consider distance without causality (beta=0.5)

21:46:42 [M] Process mining finished.

Organizational Miner

- Main idea: Which performers are executing which tasks
- Methods to mine ***roles***
  - Default mining
  - ***Doing Similar Tasks***
- Methods to mine ***teams***
  - Working together



Settings for mining Filtered repairExample.mxml (Simple f...

Organizational Miner

Method for OM

○ Default Mining

● Doing Similar Task (DST)

○ Working Together (WT)

Options for DST

○ Euclidian distance

● Correlation coefficient

○ Similarity coefficient

○ Hamming distance

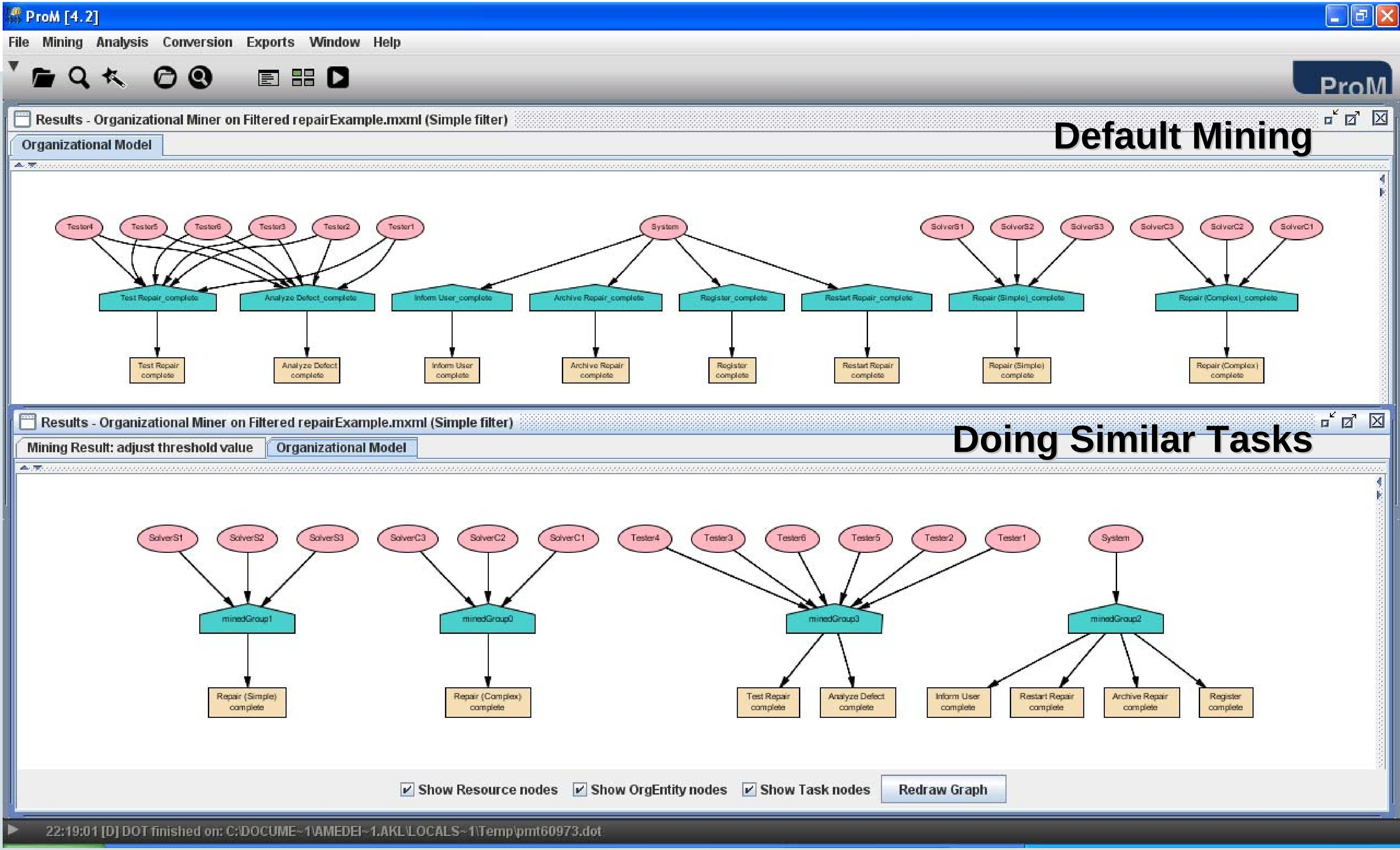
Options for WT

● Simultaneous appearance ratio

○ Consider distance with causality

○ Consider distance without causality (beta=0.5)

21:46:42 [M] Process mining finished.



Default Mining

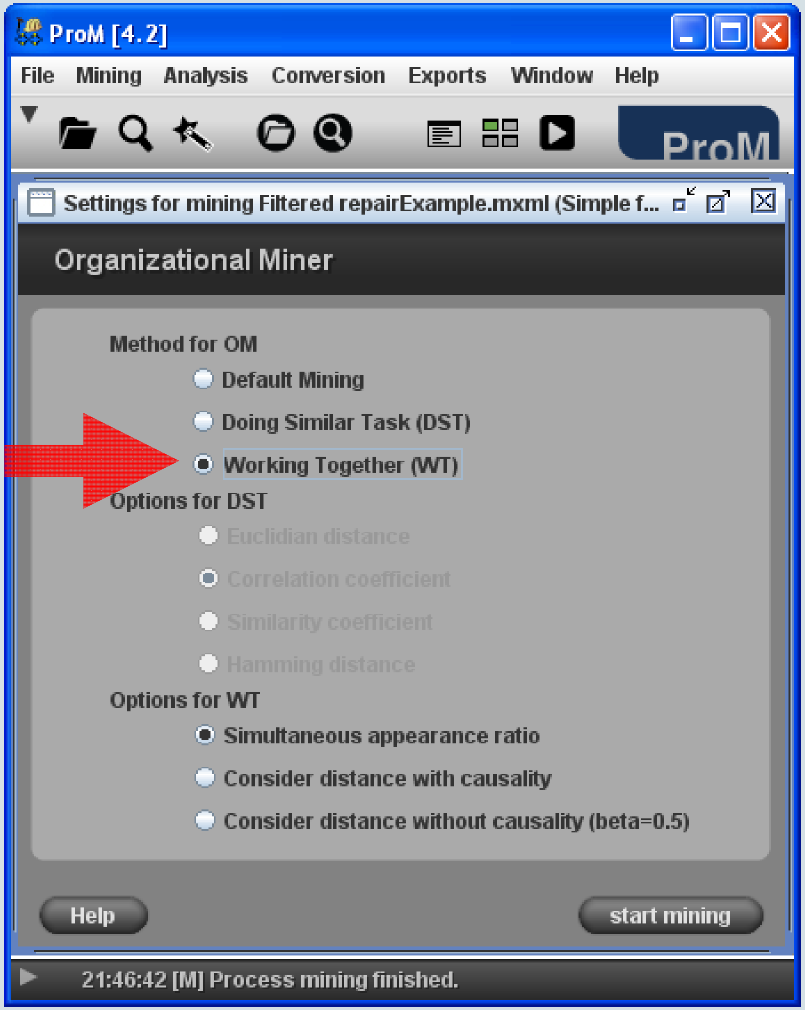
Doing Similar Tasks

☑ Show Resource nodes ☑ Show OrgEntity nodes ☑ Show Task nodes

```
22:19:01 [D] DOT finished on: C:\DOCUME~1\AMEDEI~1.AKL\LOCALS~1\Temp\pmt60973.dot
```

Organizational Miner

- Main idea: Which performers are executing which tasks
- Methods to mine ***roles***
  - Default mining
  - Doing Similar Tasks
- Methods to mine ***teams***
  - ***Working together***



Settings for mining Filtered repairExample.mxml (Simple f...

Organizational Miner

Method for OM

○ Default Mining

○ Doing Similar Task (DST)

● Working Together (WT)

Options for DST

○ Euclidian distance

○ Correlation coefficient

○ Similarity coefficient

○ Hamming distance

Options for WT

● Simultaneous appearance ratio

○ Consider distance with causality

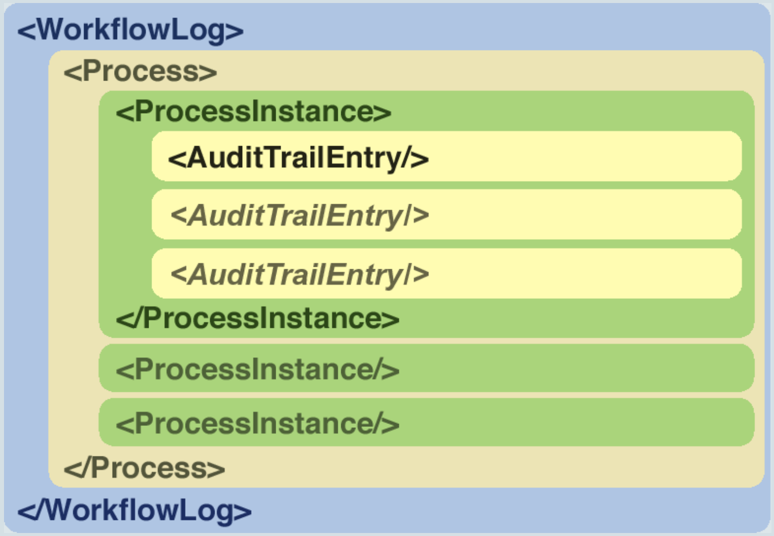
○ Consider distance without causality (beta=0.5)

21:46:42 [M] Process mining finished.

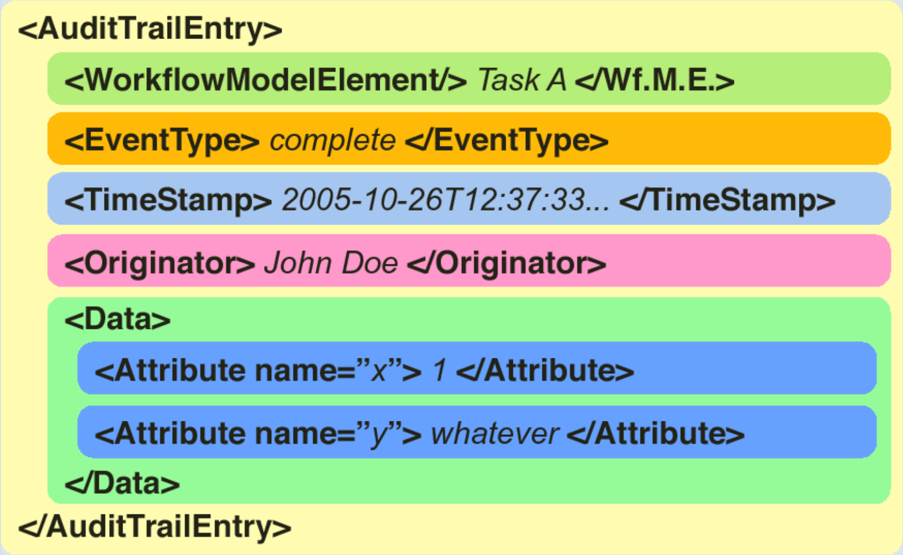
Organizational Miner

Why is the notion of process instances necessary to mine teams but unnecessary to mine roles?

Could you think of an algorithm to detect specialists/generalists for a given process? What is the main idea behind?



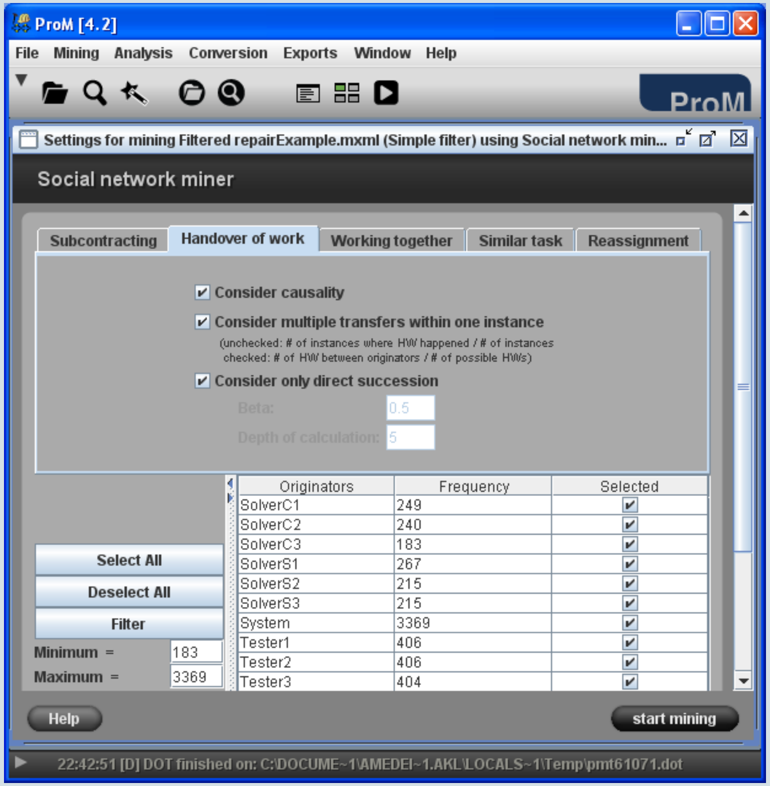
```
<WorkflowLog>
  <Process>
    <ProcessInstance>
      <AuditTrailEntry/>
      <AuditTrailEntry/>
      <AuditTrailEntry/>
    </ProcessInstance>
    <ProcessInstance/>
    <ProcessInstance/>
  </Process>
</WorkflowLog>
```



```
<AuditTrailEntry>
  <WorkflowModelElement/> Task A </Wf.M.E.>
  <EventType> complete </EventType>
  <TimeStamp> 2005-10-26T12:37:33... </TimeStamp>
  <Originator> John Doe </Originator>
  <Data>
    <Attribute name=”x”> 1 </Attribute>
    <Attribute name=”y”> whatever </Attribute>
  </Data>
</AuditTrailEntry>
```

Social Network Miner

- Aim: Monitor how individual process instances are routed between originators
- Metrics
  - – Handover of work
  - – Subcontracting
  - – Reassignment
  - – Working together
  - – Similar task



Settings for mining Filtered repairExample.mxml (Simple filter) using Social network min...

Social network miner

☑ Consider causality

☑ Consider multiple transfers within one instance
(unchecked: # of instances where HW happened / # of instances
checked: # of HW between originators / # of possible HWs)

☑ Consider only direct succession

Beta: 0.5

Depth of calculation: 5

Minimum = 183

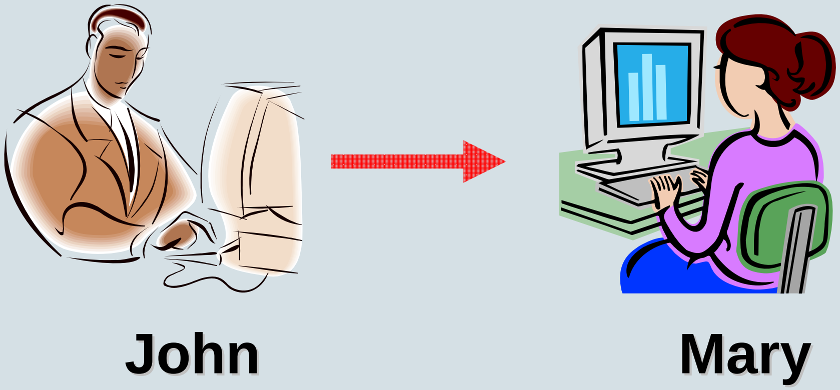
Maximum = 3369

Originators	Frequency	Selected
SolverC1	249	☑
SolverC2	240	☑
SolverC3	183	☑
SolverS1	267	☑
SolverS2	215	☑
SolverS3	215	☑
System	3369	☑
Tester1	406	☑
Tester2	406	☑
Tester3	404	☑

```
22:42:51 [D] DOT finished on: C:\DOCUME~1\AMEDEI~1.AKL\LOCALS~1\Temp\pmt61071.dot
```

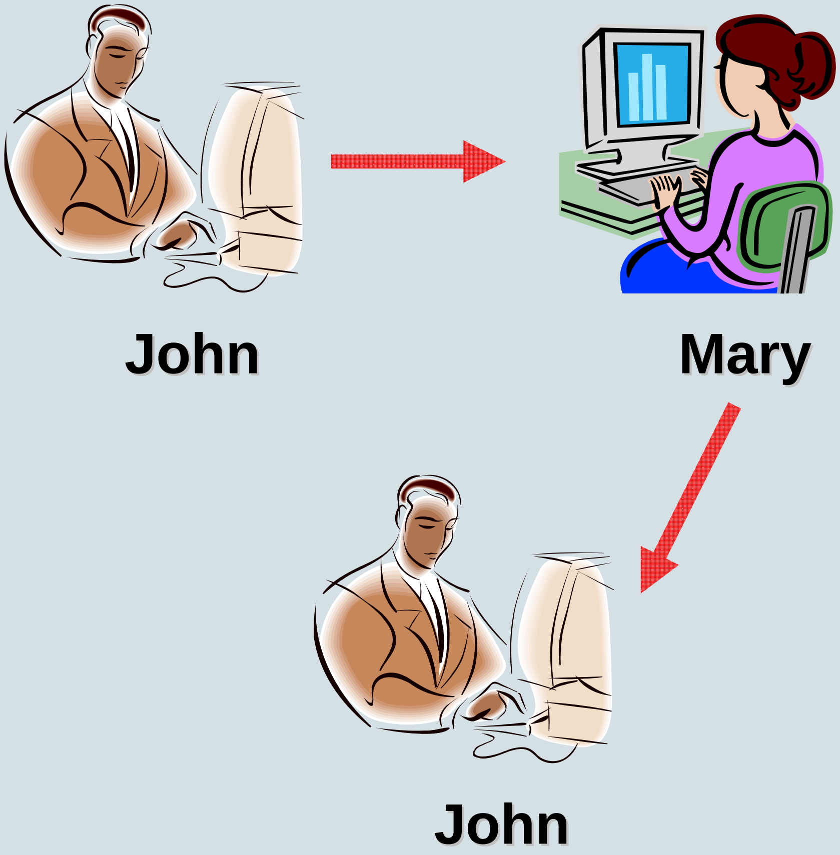
Social Network Miner

- Aim: Monitor how individual process instances are routed between originators
- Metrics
  - ***Handover of work***
  - Subcontracting
  - Reassignment
  - Working together
  - Similar task



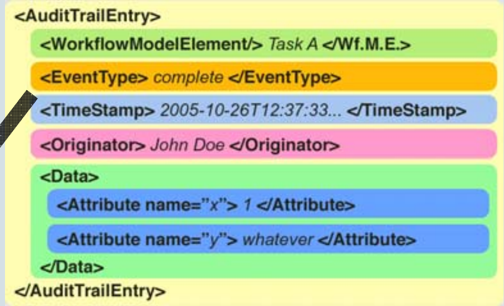
Social Network Miner

- Aim: Monitor how individual process instances are routed between originators
- Metrics
  - Handover of work
  - ***Subcontracting***
  - Reassignment
  - Working together
  - Similar task

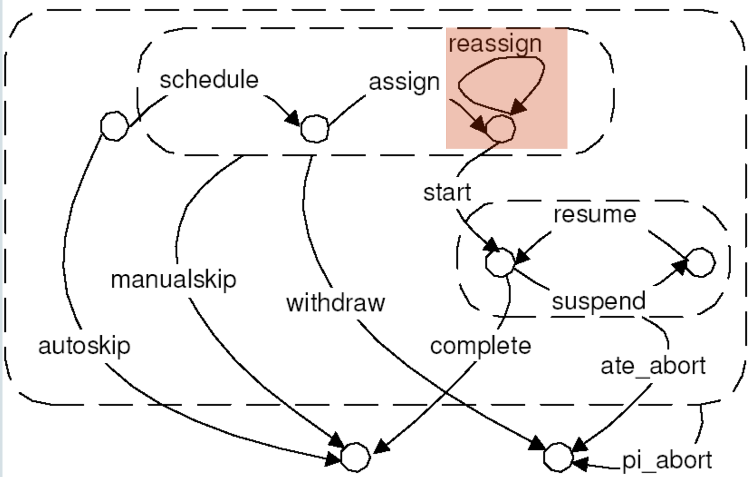


Social Network Miner

- Aim: Monitor how individual process instances are routed between originators
- Metrics
  - Handover of work
  - Subcontracting
  - ***Reassignment***
  - Working together
  - Similar task

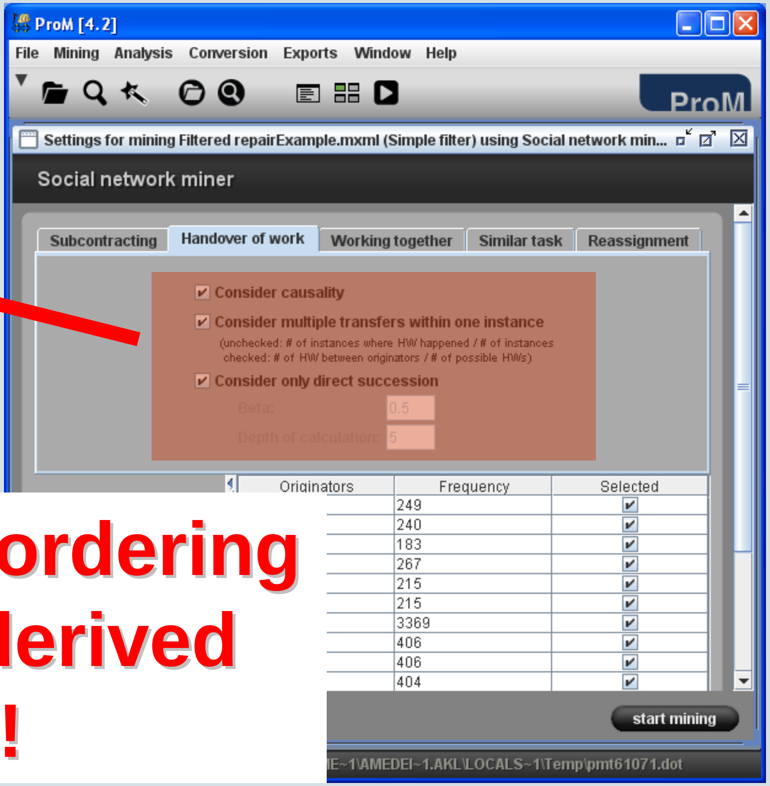


```
<AuditTrailEntry>
  <WorkflowModelElement/> Task A </Wf.M.E.>
  <EventType> complete </EventType>
  <TimeStamp> 2005-10-26T12:37:33... </TimeStamp>
  <Originator> John Doe </Originator>
  <Data>
    <Attribute name="x"> 1 </Attribute>
    <Attribute name="y"> whatever </Attribute>
  </Data>
</AuditTrailEntry>
```



Social Network Miner

- Aim: Monitor how individual process instances are routed between originators
- Metrics
  - – Handover of work
  - – Subcontracting
  - – Reassignment
  - – Working together
  - – Similar task



Settings for mining Filtered repairExample.mxml (Simple filter) using Social network min...

Social network miner

☑ Consider causality

☑ Consider multiple transfers within one instance
(unchecked: # of instances where HW happened / # of instances
checked: # of HW between originators / # of possible HWs)

☑ Consider only direct succession

Beta: 0.5

Depth of calculation: 5

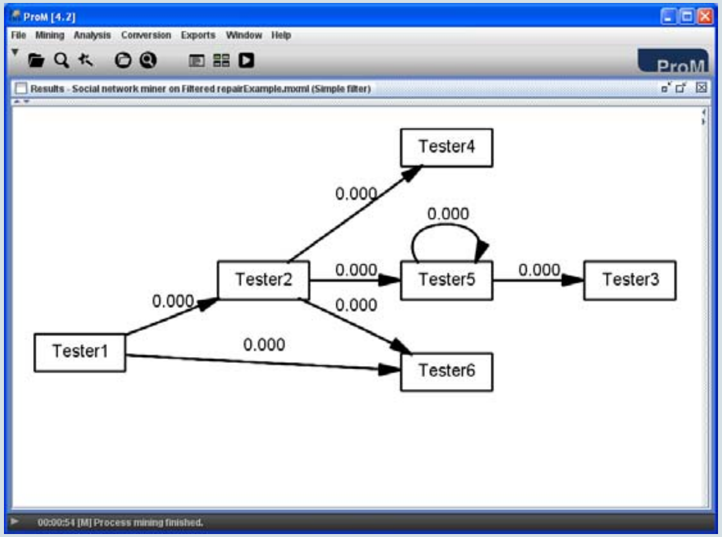
Originators	Frequency	Selected
	249	☑
	240	☑
	183	☑
	267	☑
	215	☑
	215	☑
	3369	☑
	406	☑
	406	☑
	404	☑

IE~1\AMEDEI~1.AKL\LOCALS~1\Temp\pmt61071.dot

Based on ordering relations derived from a log!

Analyze Social Network

- Better graphical view for the results of the Social Network Miner
- Includes different metrics to measure centrality of nodes
- Example: subcontracting

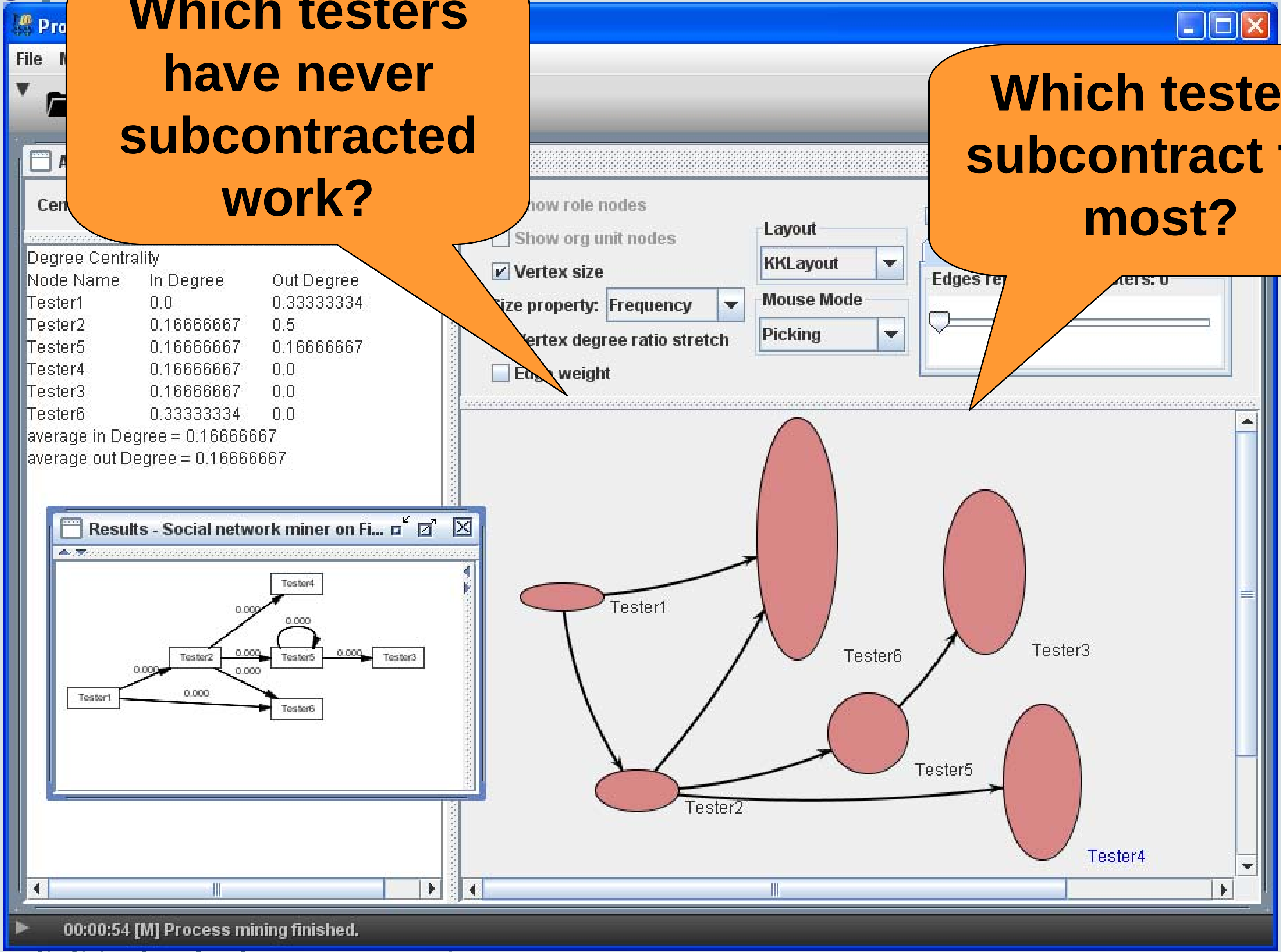


Results - Social network miner on Filtered repairExample.mxml (Simple filter)

00:00:54 [M] Process mining finished.

Which testers have never subcontracted work?

Which testers subcontract the most?



```
Degree Centrality
Node Name    In Degree     Out Degree
Tester1      0.0           0.33333334
Tester2      0.16666667    0.5
Tester5      0.16666667    0.16666667
Tester4      0.16666667    0.0
Tester3      0.16666667    0.0
Tester6      0.33333334    0.0
average in Degree = 0.16666667
average out Degree = 0.16666667
```

□ Show org unit nodes

☑ Vertex size

Size property: Frequency

Vertex degree ratio stretch

□ Edge weight

Layout: KKLayout

Mouse Mode: Picking

00:00:54 [M] Process mining finished.

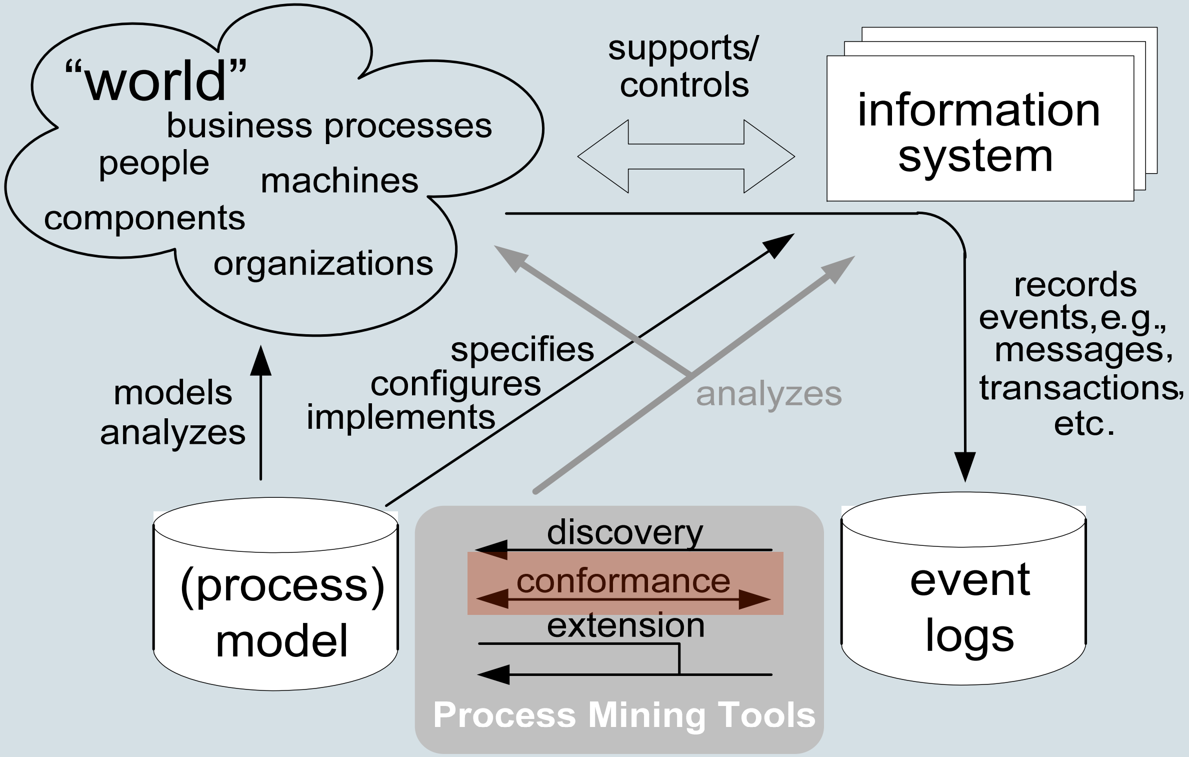
Process Mining

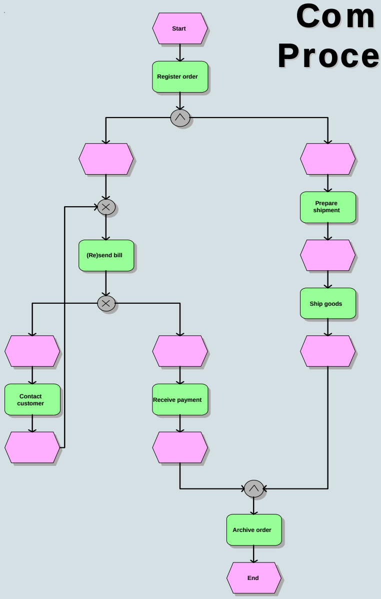
- Short Recap
- Discovery Techniques (Part 2)
  - Organizational Model
  - Social Network
- Conformance Techniques
  - Conformance Checker
  - LTL- Checker
- Summary
- Announcements

Process Mining

- Short Recap
- Discovery Techniques (Part 2)
  - Organizational Model
  - Social Network
- Conformance Techniques
  - Conformance Checker
  - LTL- Checker
- Summary
- Announcements

Types of Algorithms





Compliance Process Model

Auditing / Security

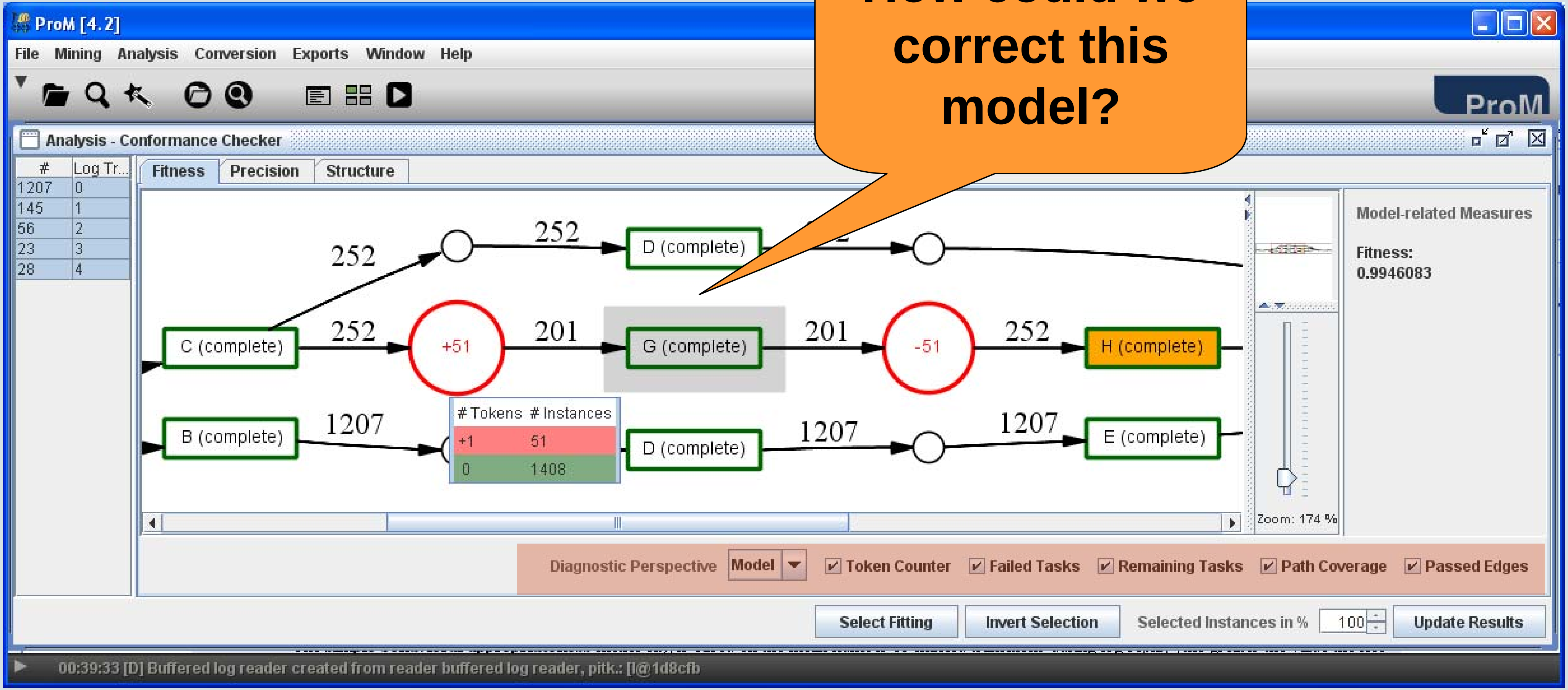
Conformance Checker

- Aim
  - Assess how much a process model matches given process instances
- Driving force
  - Replay process instances in models
- Types of diagnosis
  - Fitness
  - Structural Appropriateness
  - Behavioral Appropriateness

Fitness

- Can the model replay the log?

How could we correct this model?



ProM [4.2]
File Mining Analysis Conversion Exports Window Help
Analysis - Conformance Checker
Fitness Precision Structure
252
252
D (complete)
C (complete)
252
+51
201
G (complete)
201
-51
252
H (complete)
B (complete)
1207
D (complete)
1207
1207
E (complete)
Zoom: 174 %
Diagnostic Perspective
Model
☑ Token Counter ☑ Failed Tasks ☑ Remaining Tasks ☑ Path Coverage ☑ Passed Edges
Select Fitting
Invert Selection
Selected Instances in % 100
Update Results

#	Log Tr...
1207	0
145	1
56	2
23	3
28	4

# Tokens	# Instances
+1	51
0	1408

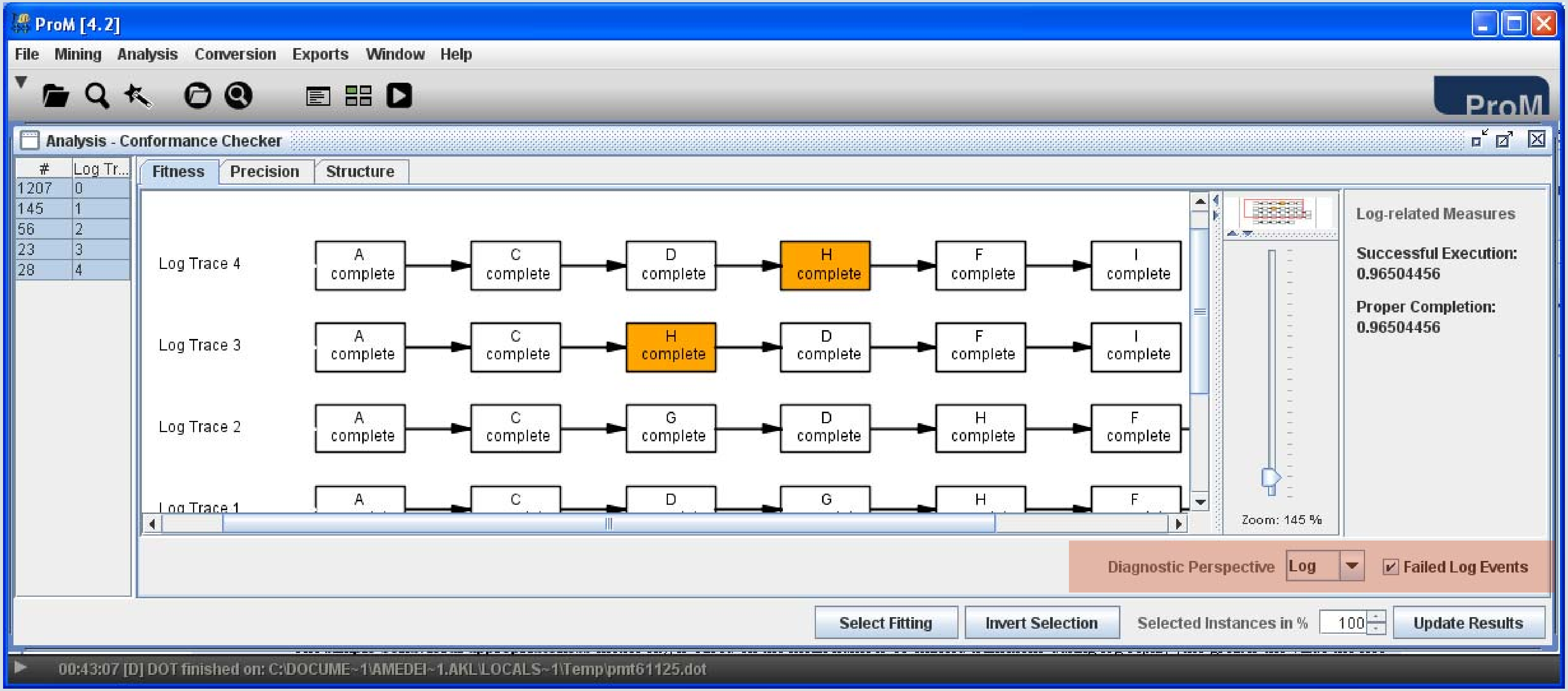
Model-related Measures

Fitness:
0.9946083

```
00:39:33 [D] Buffered log reader created from reader buffered log reader, pitk.: [I@1d8cfb
```

Fitness

- Can the model replay the log?



ProM [4.2]
File Mining Analysis Conversion Exports Window Help
ProM
Analysis - Conformance Checker
Fitness Precision Structure
Zoom: 145 %
Select Fitting
Invert Selection
Update Results

#	Log Tr...
1207	0
145	1
56	2
23	3
28	4

Log Trace 4: A complete → C complete → D complete → H complete → F complete → I complete

Log Trace 3: A complete → C complete → H complete → D complete → F complete → I complete

Log Trace 2: A complete → C complete → G complete → D complete → H complete → F complete

Log Trace 1: A → C → D → G → H → F

Log-related Measures

Successful Execution: 0.96504456

Proper Completion: 0.96504456

Diagnostic Perspective: Log

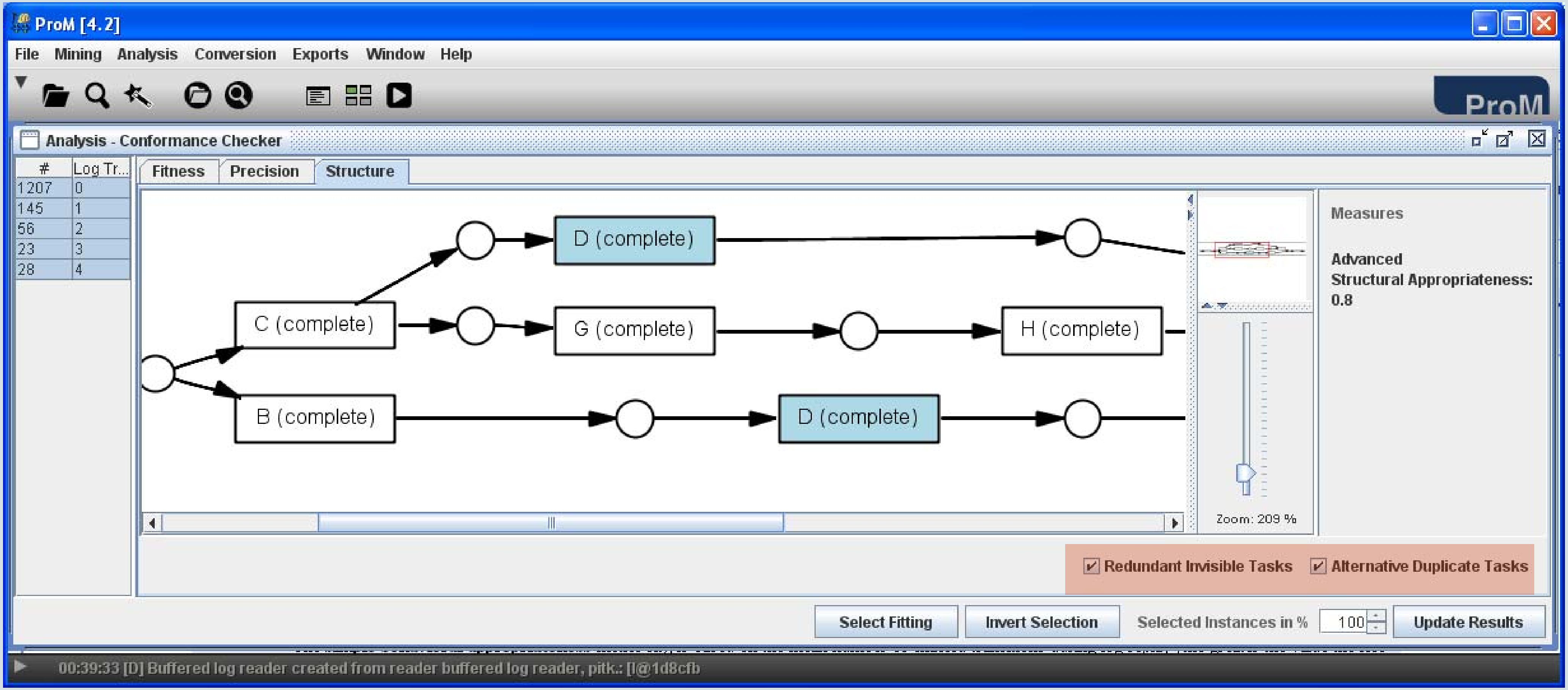
☑ Failed Log Events

Selected Instances in %: 100

```
00:43:07 [D] DOT finished on: C:\DOCUME~1\AMEDEI~1.AKL\LOCALS~1\Temp\pmt61125.dot
```

Structural Appropriateness

- Is the model overly complex?



#	Log Tr...
1207	0
145	1
56	2
23	3
28	4

Measures

Advanced Structural Appropriateness: 0.8

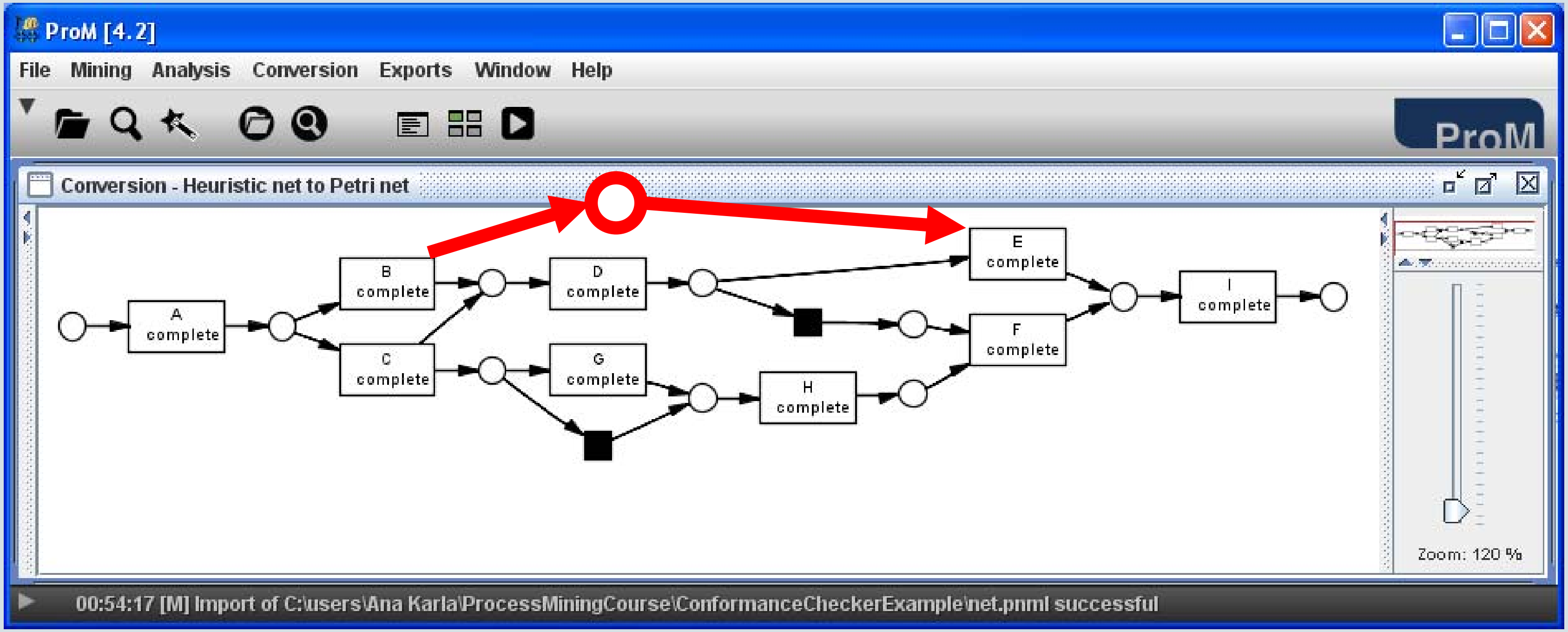
☑ Redundant Invisible Tasks ☑ Alternative Duplicate Tasks

Selected Instances in % 100

```
00:39:33 [D] Buffered log reader created from reader buffered log reader, pitk.: [I@1d8cfb
```

Behavioral Appropriateness

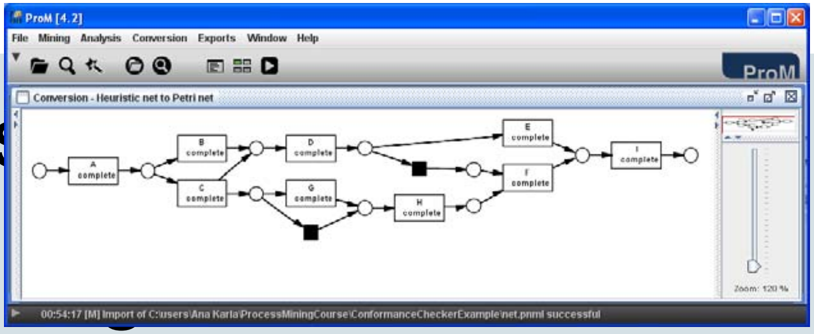
- Another example



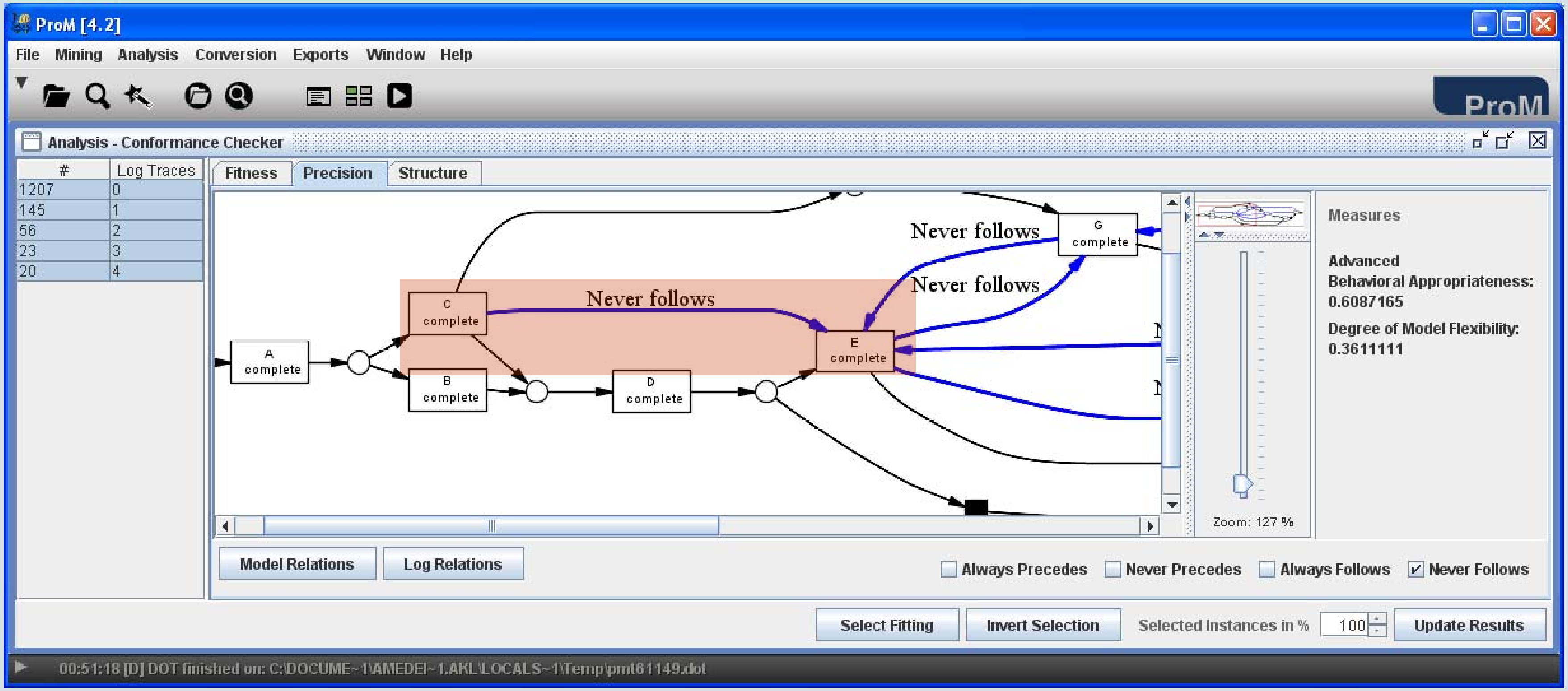
00:54:17 [M] Import of C:\users\Ana Karla\ProcessMiningCourse\ConformanceCheckerExample\net.pnml successful

Behavioral Appropriatenes

- Is the model precise en



00:54:17 [M] Import of C:\users\Ana Karla\ProcessMiningCourse\ConformanceCheckerExample\net.pnml successful



#	Log Traces
1207	0
145	1
56	2
23	3
28	4

Measures

Advanced Behavioral Appropriateness: 0.6087165

Degree of Model Flexibility: 0.3611111

□ Always Precedes □ Never Precedes □ Always Follows ☑ Never Follows

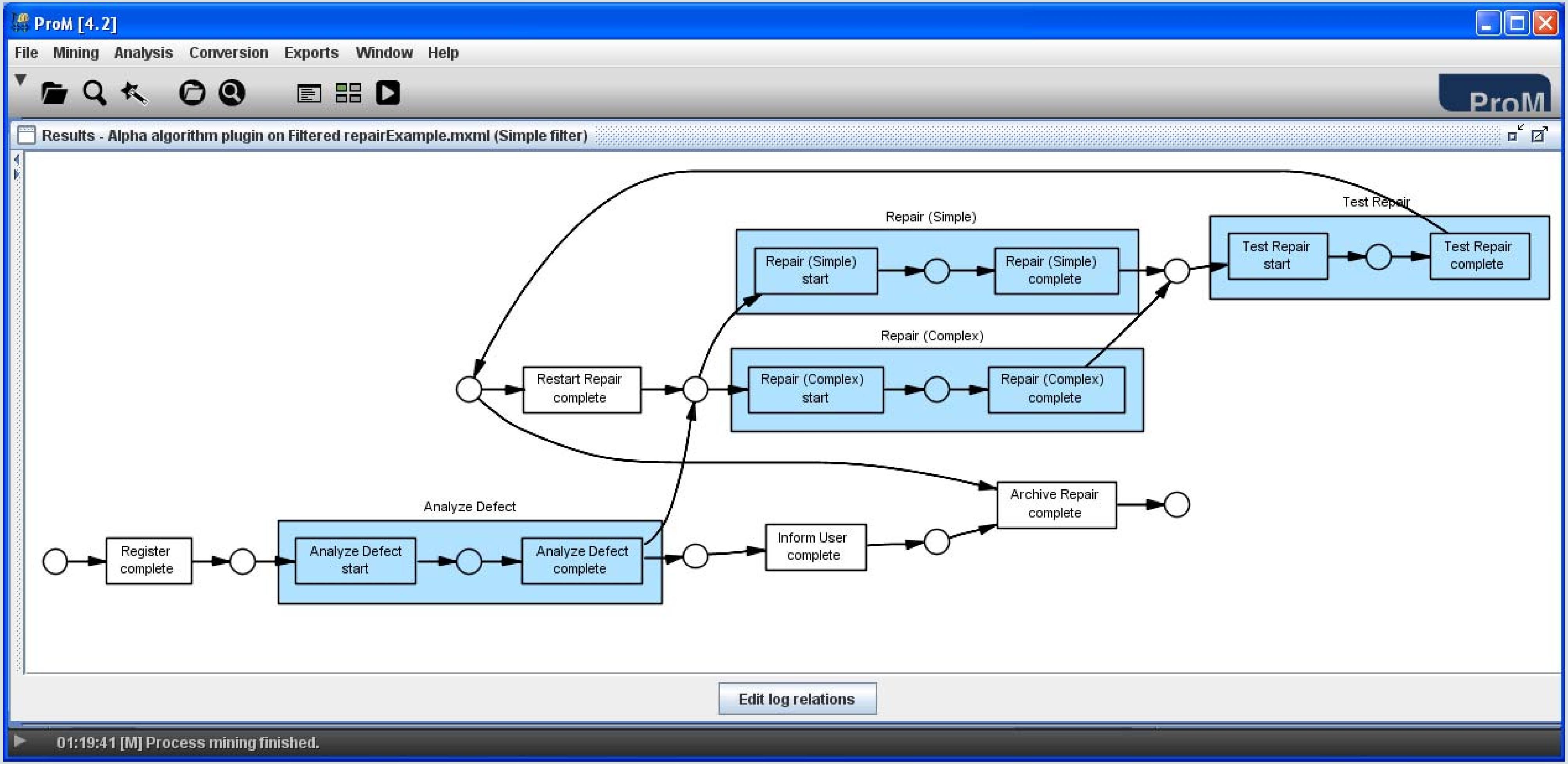
Selected Instances in % 100

00:51:18 [D] DOT finished on: C:\DOCUME~1\AMEDEI~1.AKL\LOCALS~1\Temp\pmt61149.dot

LTL Checker

- Aim
  - Verify if process instances fulfill certain properties
- Driving force
  - Specification of properties in a language based on **L**inear **T**emporal **L**ogics
- Example
  - Four-eyes principle

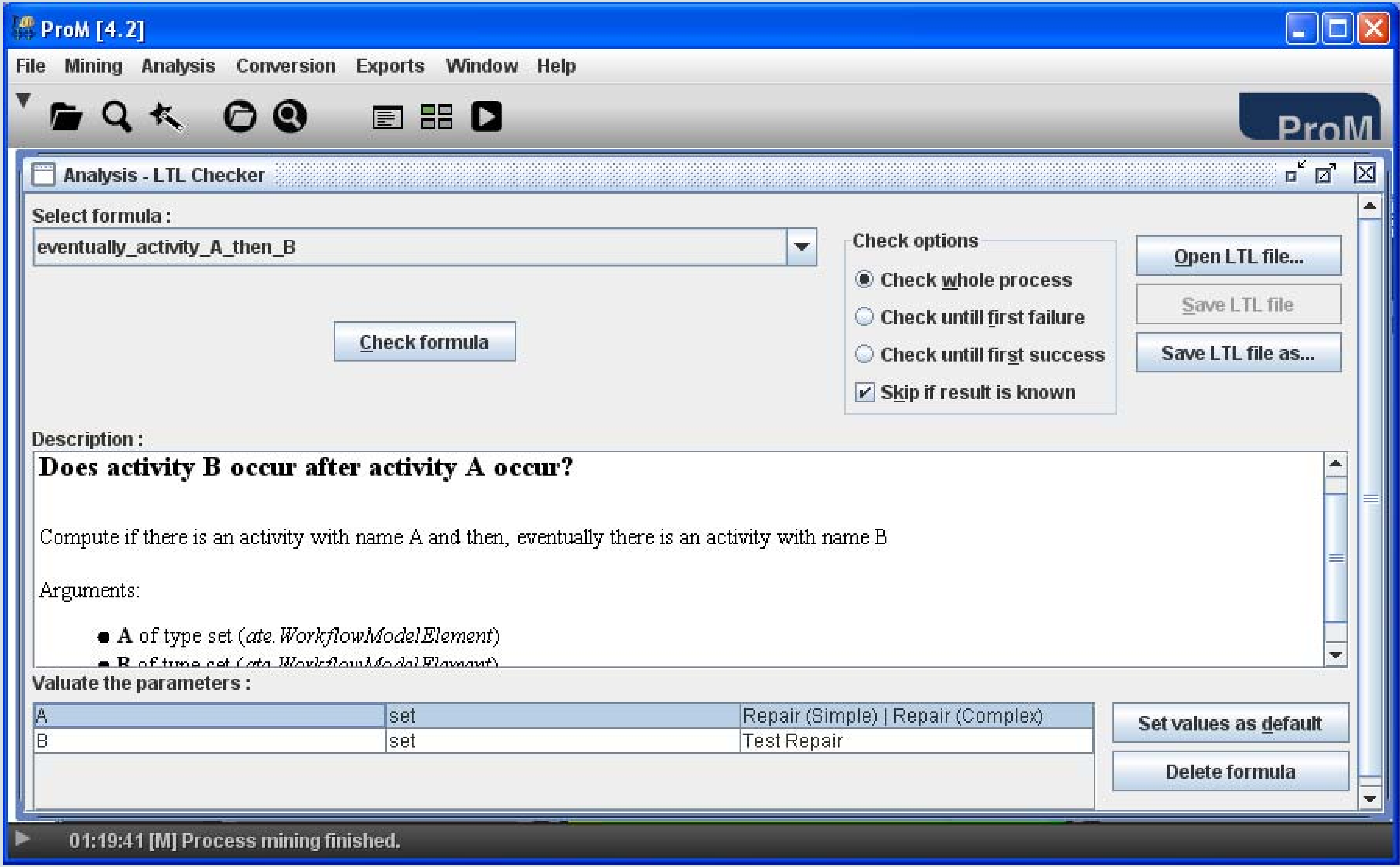
LTL Checker - Example



Results - Alpha algorithm plugin on Filtered repairExample.mxml (Simple filter)

01:19:41 [M] Process mining finished.

LTL Checker - Example



Select formula : eventually_activity_A_then_B

Check options

- ● Check whole process
- ○ Check untill first failure
- ○ Check untill first success
- ☑ Skip if result is known

Description :

Does activity B occur after activity A occur?

Compute if there is an activity with name A and then, eventually there is an activity with name B

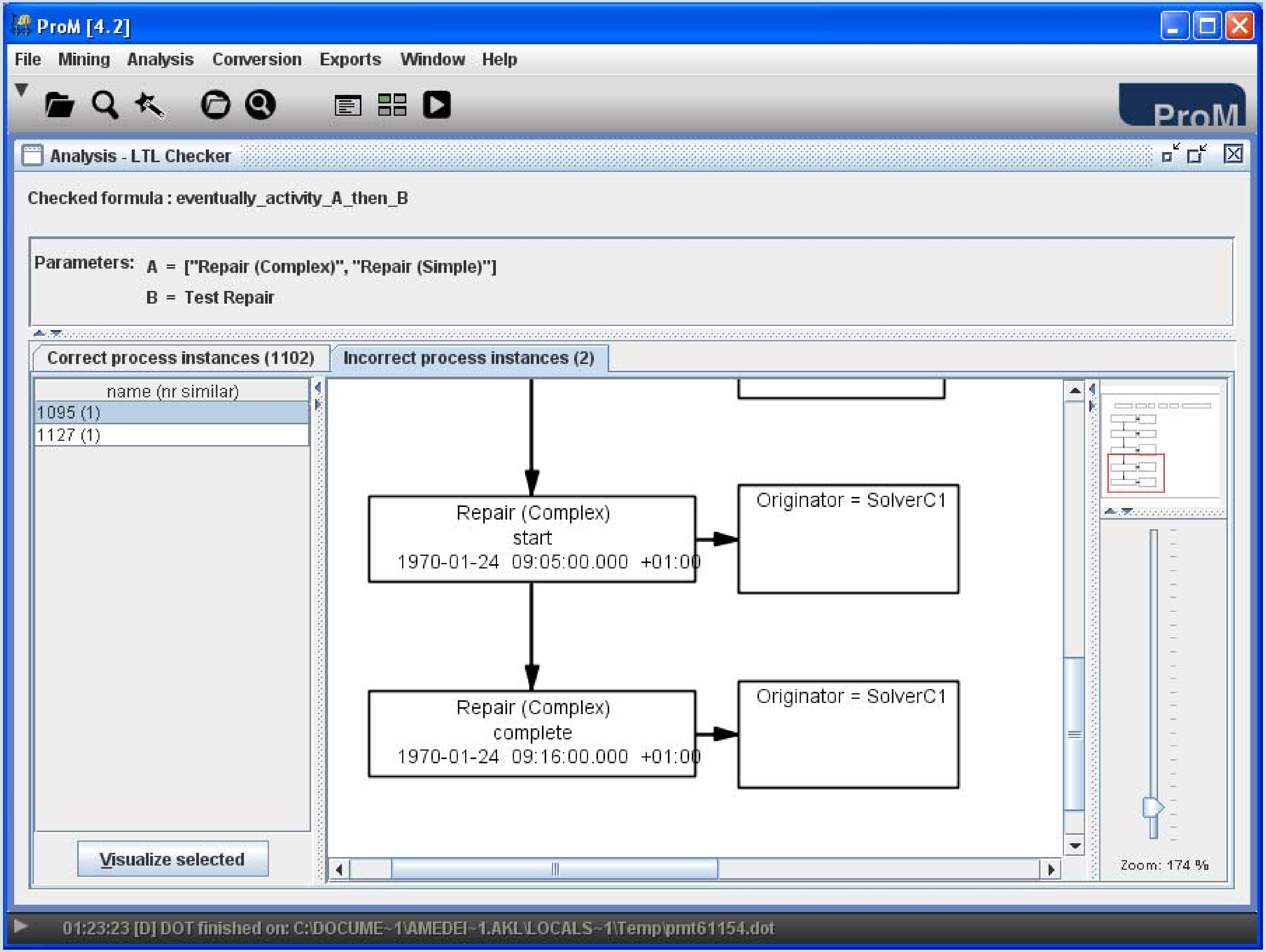
Arguments:

- **A** of type set (*ate.WorkflowModelElement*)
- B of type set (*ate.WorkflowModelElement*)

Valuate the parameters :

A	set	Repair (Simple) | Repair (Complex)
B	set	Test Repair

01:19:41 [M] Process mining finished.

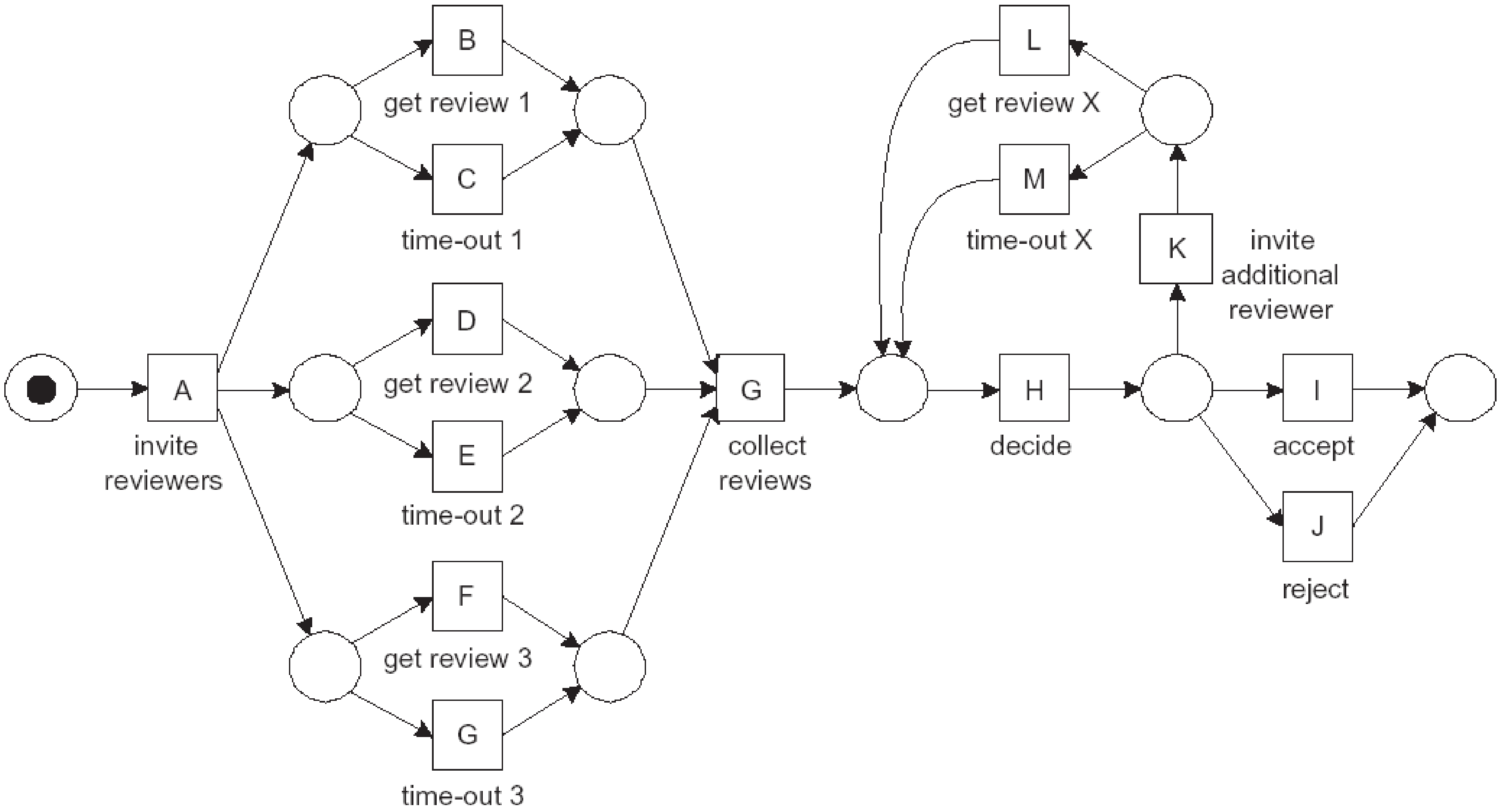


Checked formula : eventually_activity_A_then_B

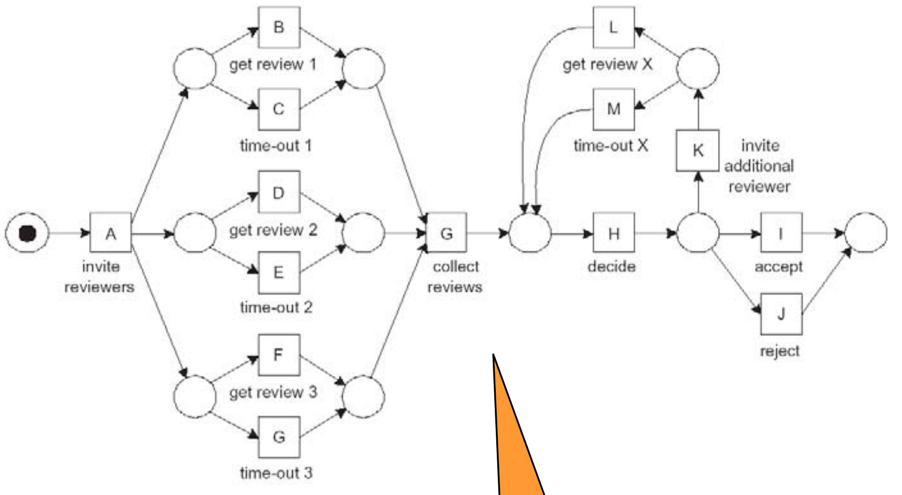
Parameters: A = ["Repair (Complex)", "Repair (Simple)"]
B = Test Repair

01:23:23 [D] DOT finished on: C:\DOCUME~1\AMEDEI~1.AKL\LOCALS~1\Temp\pmt61154.dot

LTL – Defining Formulae



```
set ate.WorkflowModelElement;
set ate.Originator;
set ate.EventType;

date ate.Timestamp := "yyyy-MM-dd";
string ate.result;
string pi.title;

rename ate.Originator as person;
rename ate.Timestamp as timestamp;
rename ate.WorkflowModelElement as activity;

formula accept_or_reject_but_not_both() :=
(<>(activity == "accept") <-> !(<>(activity == "reject")));

formula action_follows_decision() :=
[]( (activity == "decide" -> _O( ((activity == "accept" \/
activity == "reject") \/ activity == "invite additional reviewer") )));

subformula execute( p : person, a : activity ) :=
<> ( (activity == a  /\  person == p ) ) ;

formula not_the_same_reviewer() :=
forall[p:person |
(((!(execute(p,"get review 1")) \/ !(execute(p,"get review 2"))) /\
(!(execute(p,"get review 1")) \/ !(execute(p,"get review 3")))) /\
(!(execute(p,"get review 2")) \/ !(execute(p,"get review 3")))) ];
```



Can you define a formula to verify the four-eyes principle?

Summary

What are the three most important things you’ve learned today?

Summary

- Organizational mining plug-ins can discover
  - Roles/Teams in organizations
  - Social networks for originators
- Some metrics of social networks are based on ordering relations (e.g., the ordering relations used by the Alpha algorithm)
- Conformance Checker assesses how much a process model matches process instances
- LTL Checker uses logics to verify properties in event logs

Announcements

- Next lecture
  - – Invited talk **Futura Technology**
    - Start-up company in the process mining area
    - Implemented the process mining component of the BPM Suite recently release by Pallas Athena (see press release at “news” in www.processmining.org)
- Course Material
  - – See version 2 of Study Guide (posted on 18/2/2008)
- Assignments
  - – If necessary, use tutorial to get familiar with the ProM tool