

Discovering and Exploring State-based Models for Multi-perspective Processes

M.L. van Eck,
N. Sidorova,
W.M.P. van der Aalst



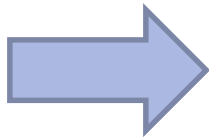
Technische Universiteit
Eindhoven
University of Technology

Where innovation starts

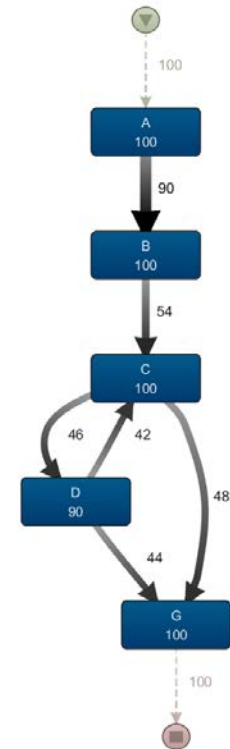
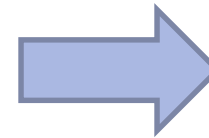
Process Discovery



Event Log



Process Discovery
Algorithm



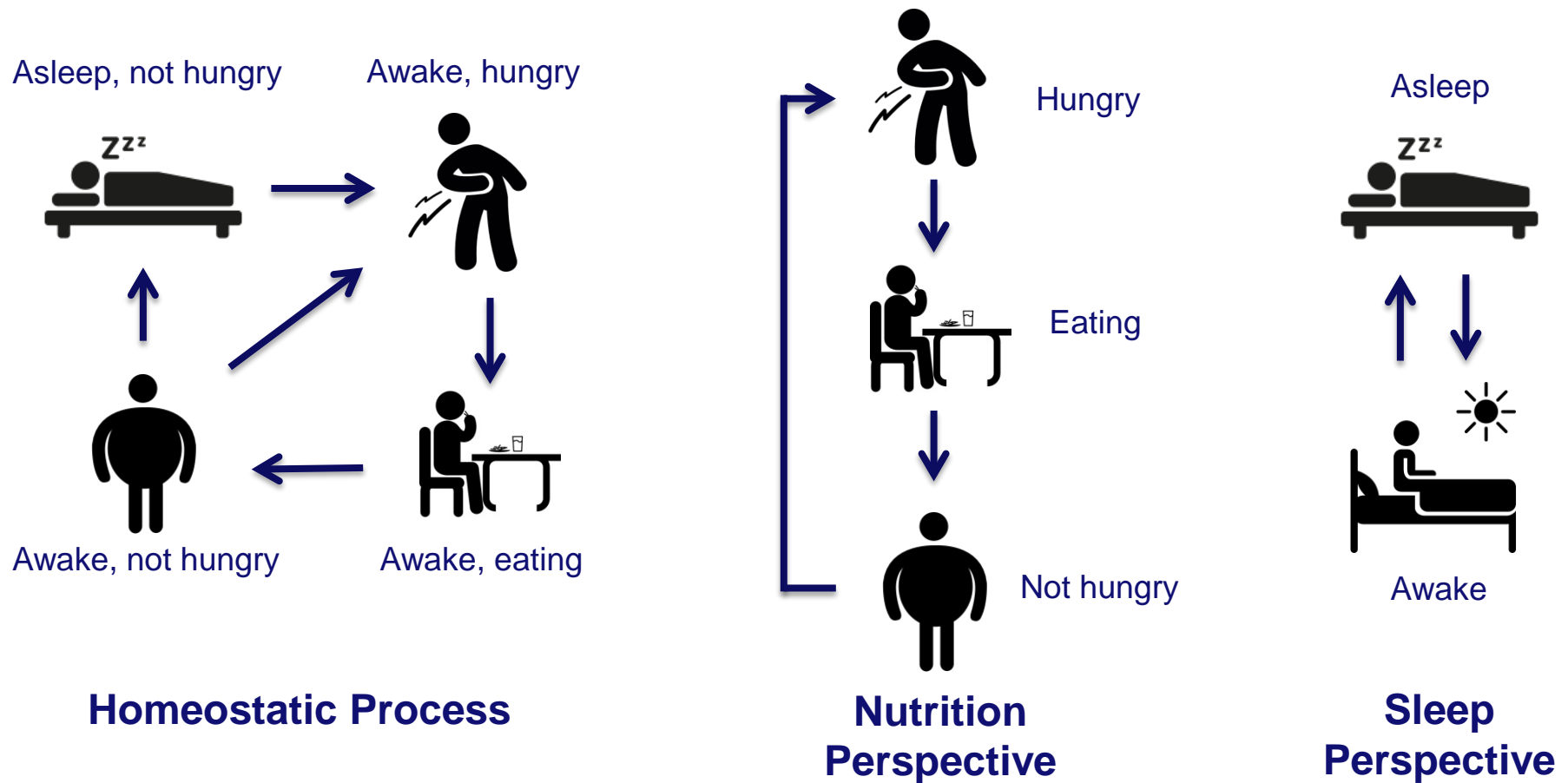
Process Model



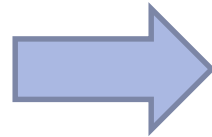
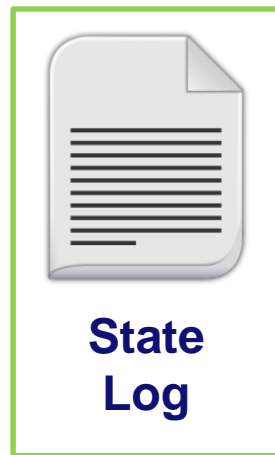
State-based Processes

Patient ID	Time	Patient Status
1	2016-03-01	Registered
1	2016-03-04	Diagnosed
1	2016-03-07	In Treatment
2	2016-03-03	Registered
2	2016-03-05	Diagnosed
3	2016-04-01	Diagnosed
...

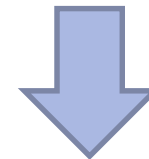
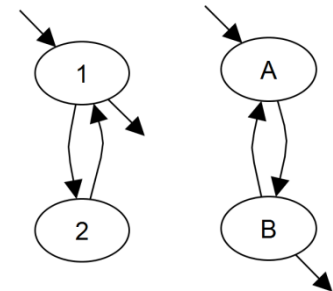
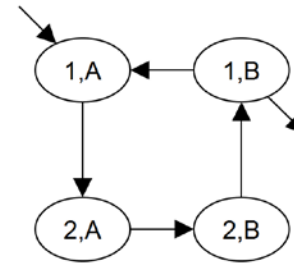
Multi-perspective State-based Processes



State-based Process Discovery

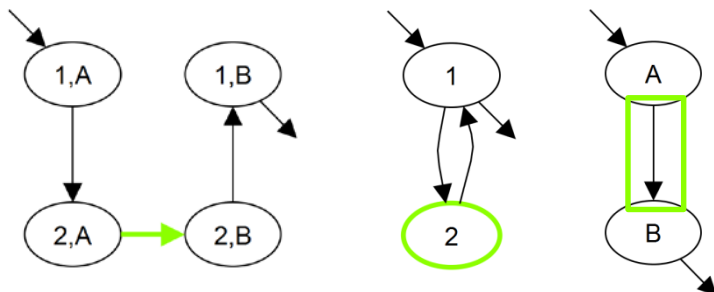
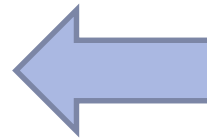


1. Discovering Multi-perspective Models

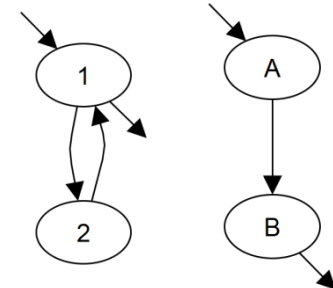
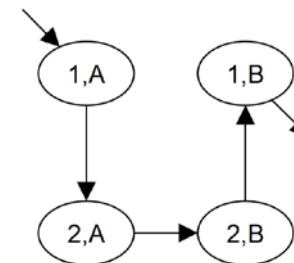


2. Creating Simplified Views

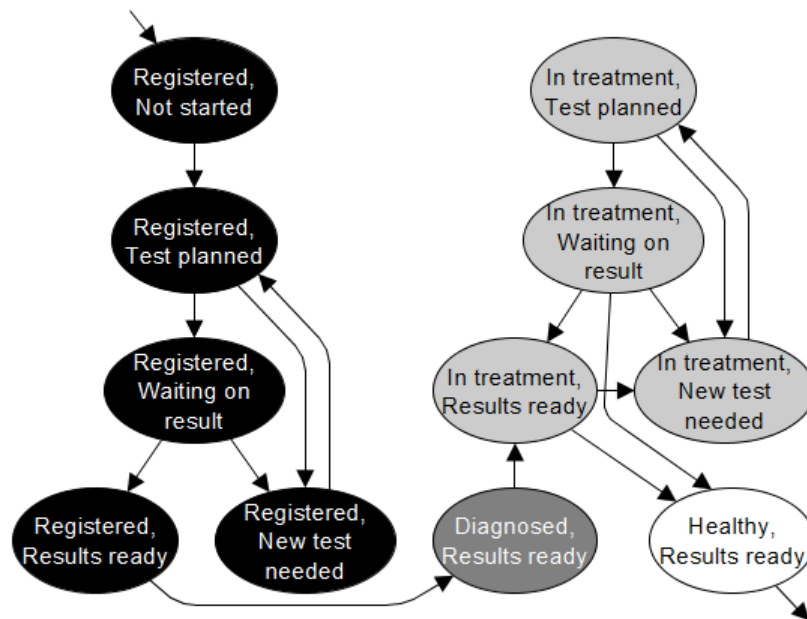
3. Computing Behavioural Relations



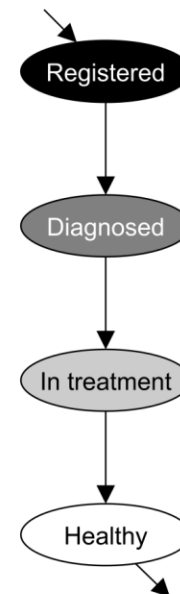
Conf: 0.4
Lift: 2.0



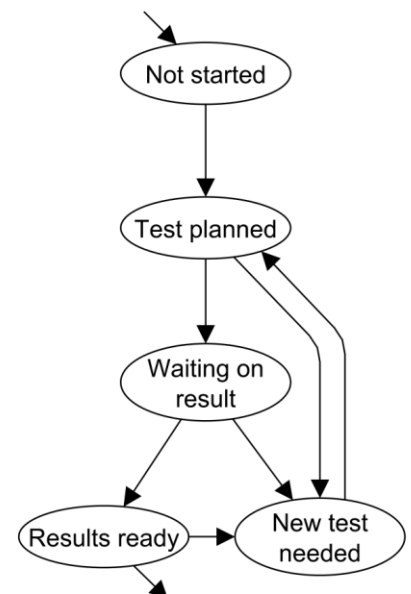
Running Example



Healthcare Process



Patient Perspective



Lab Testing Perspective



State Log

Case ID	Time	Patient State	Test State
1	2016-03-01	Registered	Not started
...



State Log

Case ID	Time	Patient State	Test State
1	2016-03-01	Registered	Not started
1	2016-03-02	Registered	Test planned
...



State Log

Case ID	Time	Patient State	Test State
1	2016-03-01	Registered	Not started
1	2016-03-02	Registered	Test planned
1	2016-03-04	Registered	Waiting on results
1	2016-03-09	Registered	Results ready
1	2016-03-10	Diagnosed	Results ready
1	2016-03-11	In treatment	Results ready
1	2016-03-21	Healthy	Results ready
2	2016-03-01	Registered	Not started
...



State Log

Case ID	Time	Patient State	Test State
1	2016-03-01	Registered	Not started
1	2016-03-02		Test planned
1	2016-03-04		Waiting on results
1	2016-03-09		Results ready
1	2016-03-10	Diagnosed	Results ready
1	2016-03-11	In treatment	Results ready
1	2016-03-21	Healthy	Results ready
2	2016-03-01	Registered	Not started
...



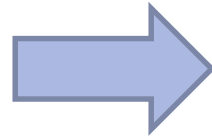
State Log

Case ID	Time	Patient State	Test State
1	2016-03-01	Registered	Not started
1	2016-03-02	Registered	Test planned
1	2016-03-04	Registered	Waiting on results
1	2016-03-09	Registered	Results ready
1	2016-03-10	Diagnosed	
1	2016-03-11	In treatment	
1	2016-03-21	Healthy	
2	2016-03-01	Registered	Not started
...

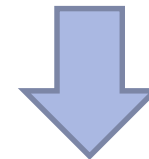
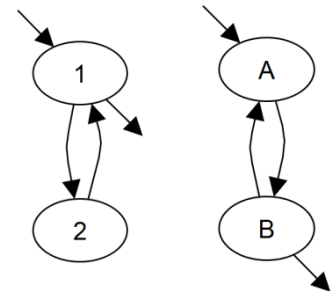
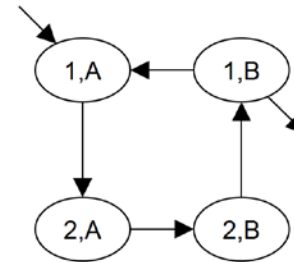
State-based Process Discovery



**State
Log**

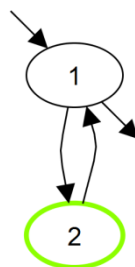
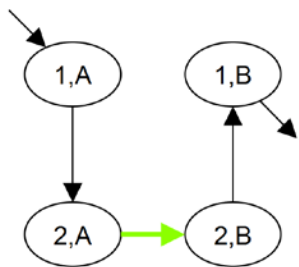
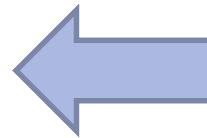


**1. Discovering
Multi-perspective
Models**

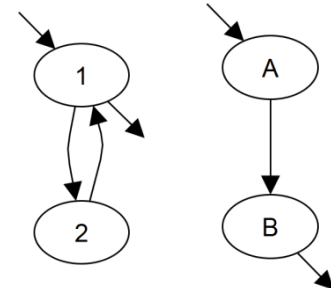
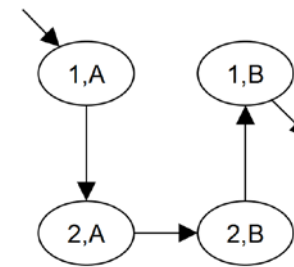
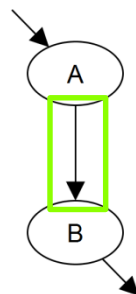


**2. Creating
Simplified
Views**

**3. Computing
Behavioural
Relations**



**Conf: 0.4
Lift: 2.0**





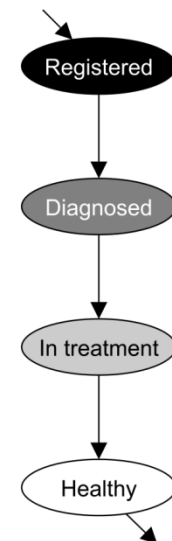
Discovering Multi-Perspective Models

- **Input: State Log**
 - Log for the entire process
 - 1 per perspective
- **Output: State Machine (SM)**
 - SM for the entire process
 - 1 SM per perspective
- **Assumptions:**
 - Perspectives and their possible states are known
 - At each point in time, the state of all perspectives is known

Discovering Multi-Perspective Models

- **Trivial discovery approach:**
 - **Create a node for each state**
 - **Draw an arc if one state changes into another**

Case ID	Time	Patient State
1	2016-03-01	Registered
1	2016-03-02	
1	2016-03-04	
1	2016-03-09	
1	2016-03-10	Diagnosed
1	2016-03-11	In treatment
1	2016-03-21	Healthy
2	2016-03-01	Registered
...

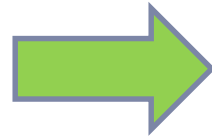


Patient State

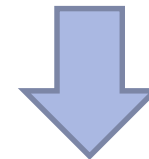
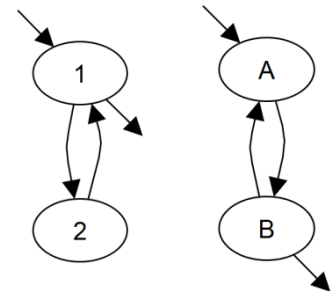
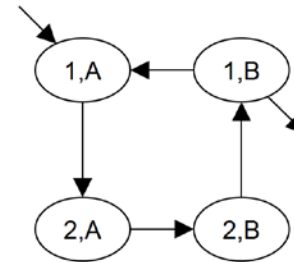
State-based Process Discovery



**State
Log**

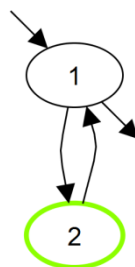
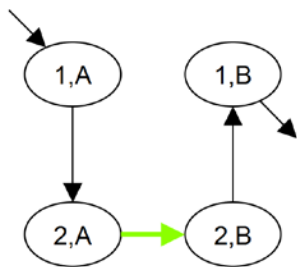
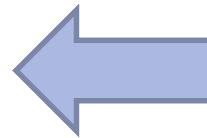


**1. Discovering
Multi-perspective
Models**

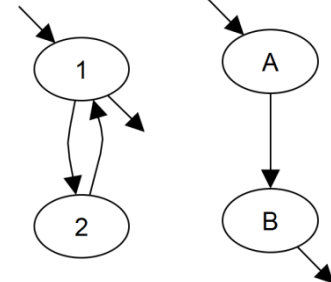
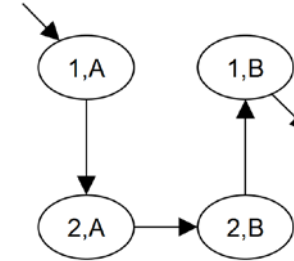
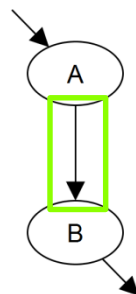


**2. Creating
Simplified
Views**

**3. Computing
Behavioural
Relations**



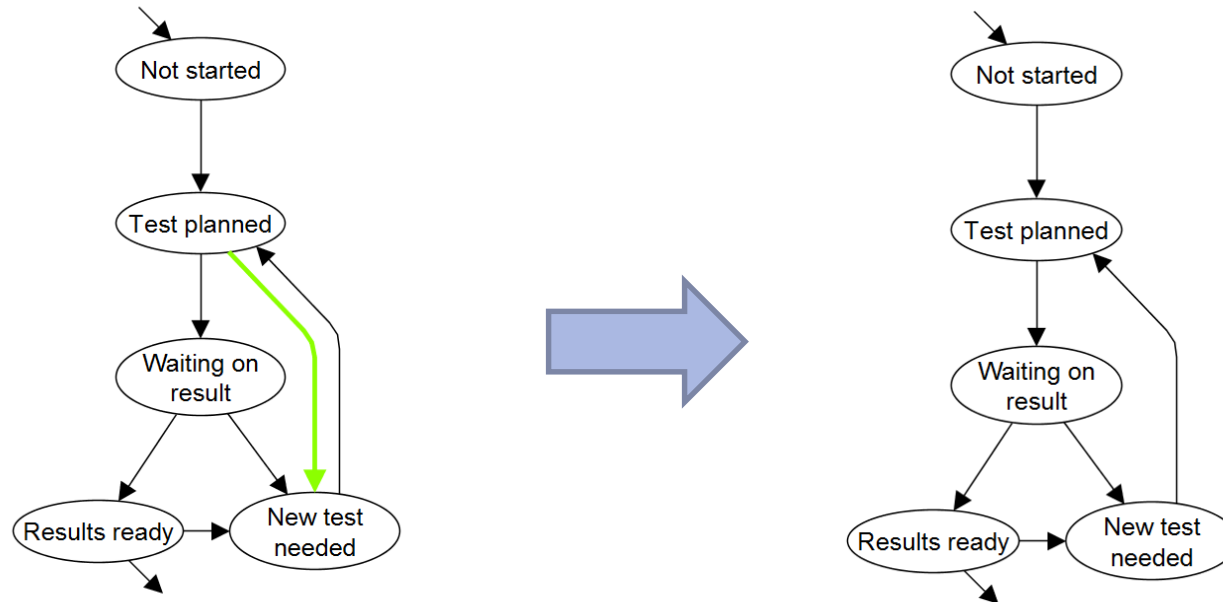
**Conf: 0.4
Lift: 2.0**



Creating Views

- **State Machines can be complex, with many states and transitions**
- **Focus analysis on part of process**
- **Three operations to create simplified views:**
 - **Removing a transition**
 - **Abstracting from a state**
 - **Aggregating two states**

Removing a Transition





Abstracting From a State

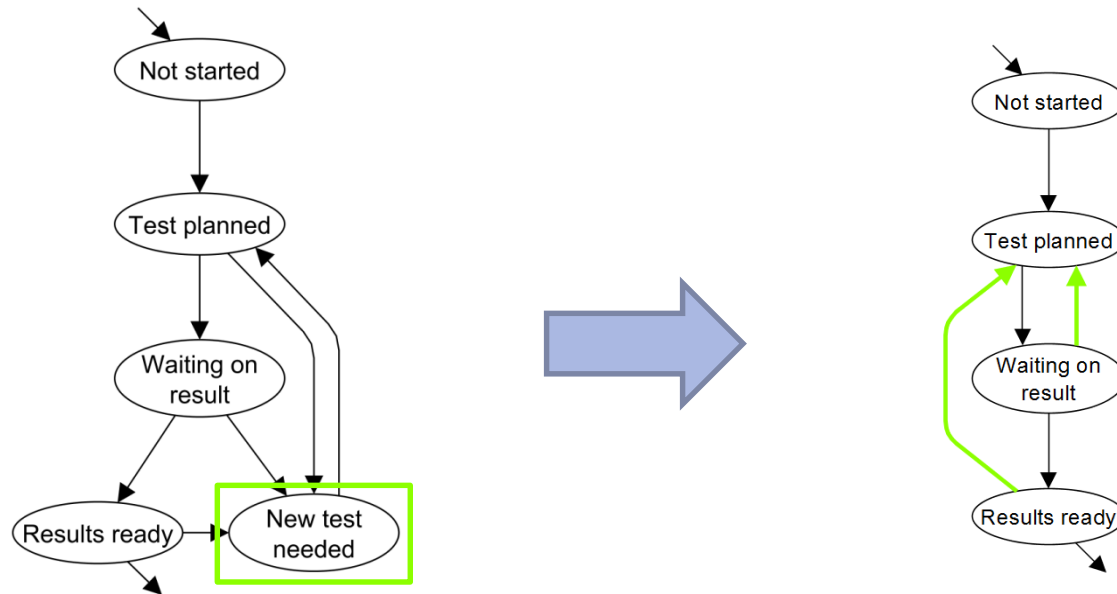
Case ID	Time	Test State
1	2016-03-01	Not started
1	2016-03-02	Test planned
1	2016-03-04	Waiting on results
1	2016-03-09	New test needed
1	2016-03-09	Test planned
1	2016-03-11	Waiting on results
1	2016-03-21	Results ready
1	2016-03-22	New test needed
1	2016-03-22	Test planned
...



Abstracting From a State

Case ID	Time	Test State
1	2016-03-01	Not started
1	2016-03-02	Test planned
1	2016-03-04	Waiting on results
1	2016-03-09	New test needed
1	2016-03-09	Test planned
1	2016-03-11	Waiting on results
1	2016-03-21	Results ready
1	2016-03-22	New test needed
1	2016-03-22	Test planned
...

Abstracting From a State





Aggregating Two States

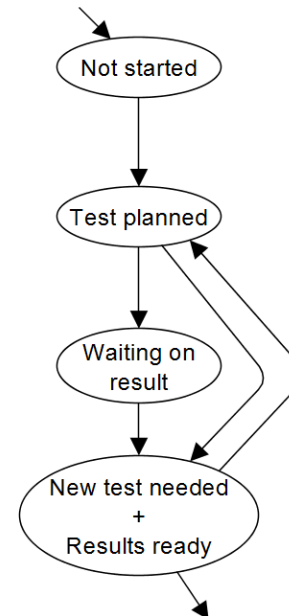
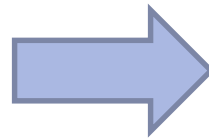
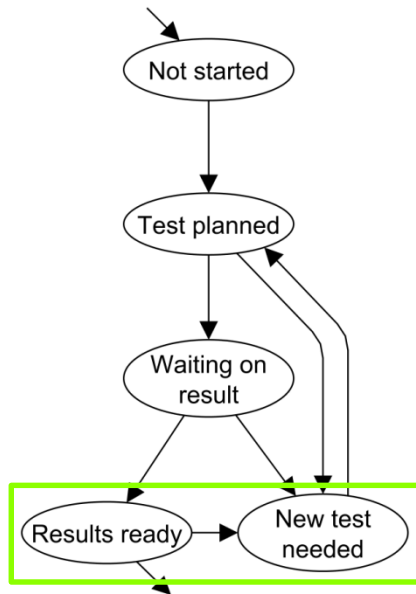
Case ID	Time	Test State
1	2016-03-01	Not started
1	2016-03-02	Test planned
1	2016-03-04	Waiting on results
1	2016-03-09	New test needed
1	2016-03-09	Test planned
1	2016-03-11	Waiting on results
1	2016-03-21	Results ready
1	2016-03-22	New test needed
1	2016-03-22	Test planned
...



Aggregating Two States

Case ID	Time	Test State
1	2016-03-01	Not started
1	2016-03-02	Test planned
1	2016-03-04	Waiting on results
1	2016-03-09	Results ready + New test needed
1	2016-03-09	Test planned
1	2016-03-11	Waiting on results
1	2016-03-21	Results ready + New test needed
1	2016-03-22	Test planned
...

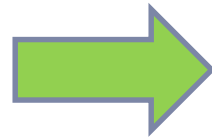
Aggregating Two States



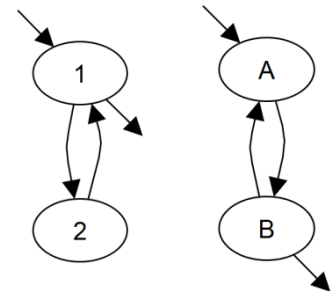
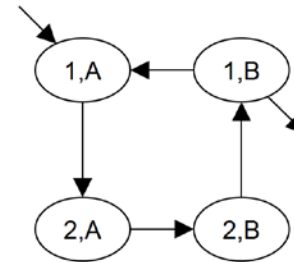
State-based Process Discovery



**State
Log**

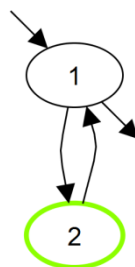
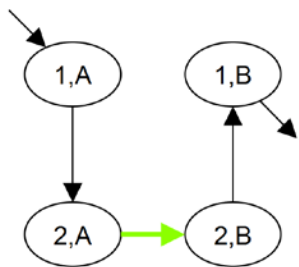
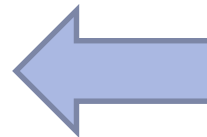


**1. Discovering
Multi-perspective
Models**

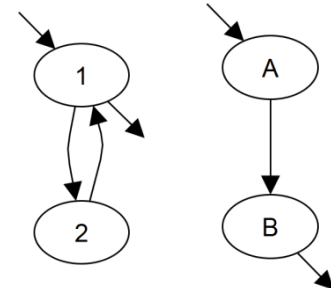
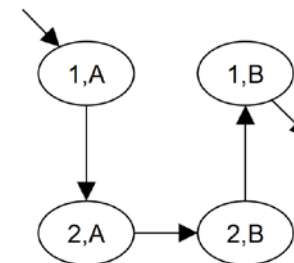
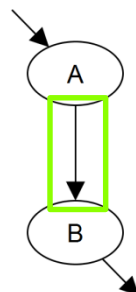


**2. Creating
Simplified
Views**

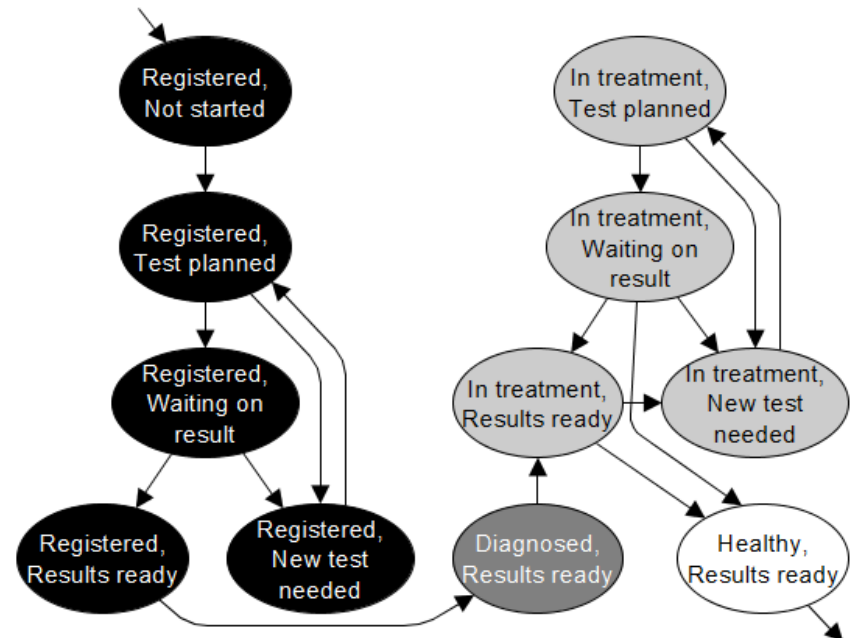
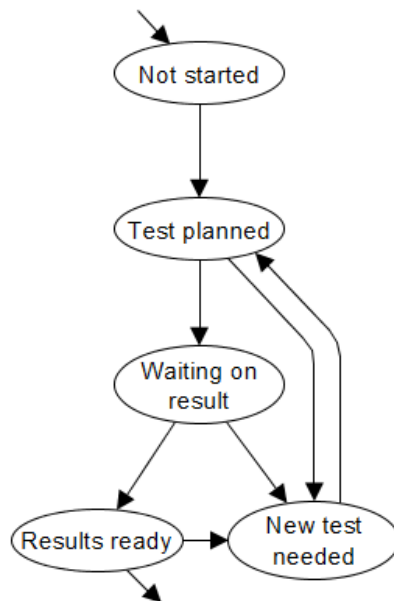
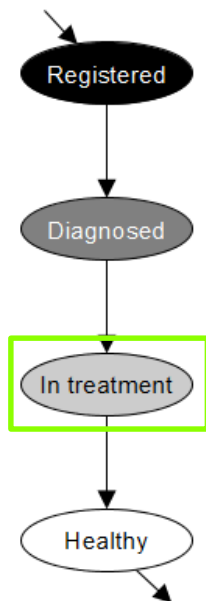
**3. Computing
Behavioural
Relations**



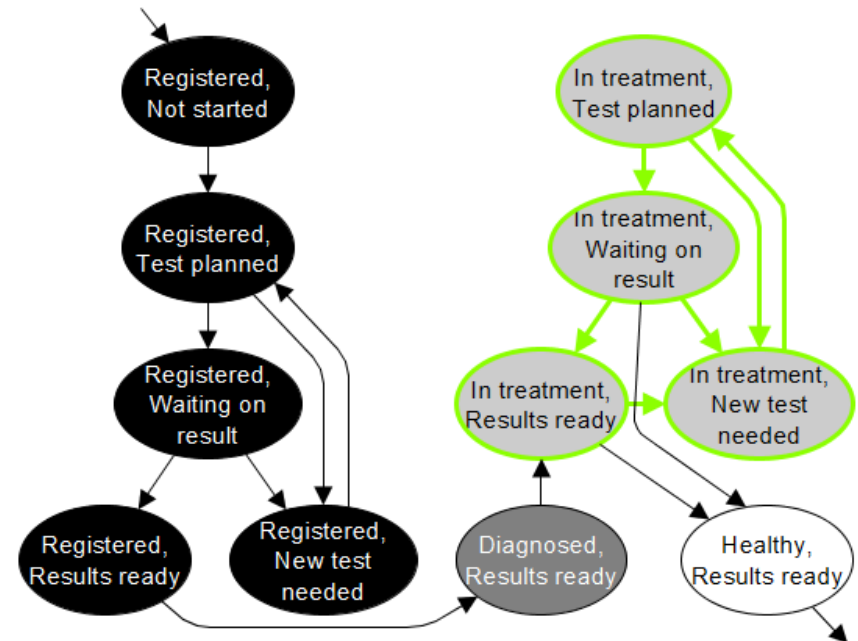
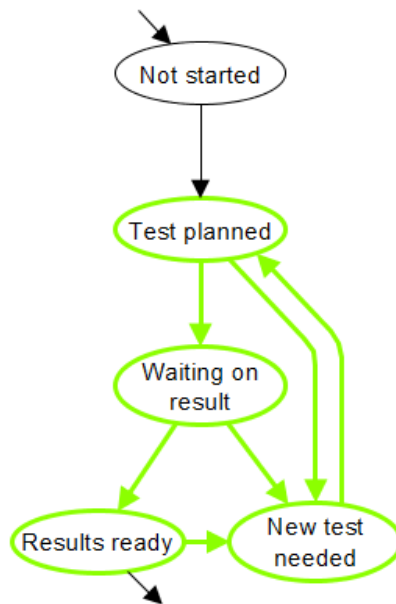
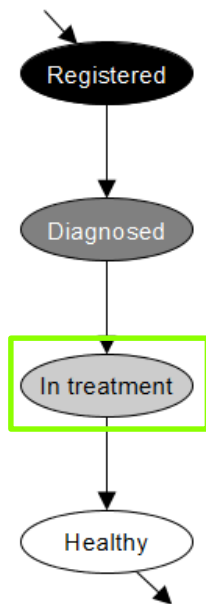
**Conf: 0.4
Lift: 2.0**



Behavioural Relations



Behavioural Relations



Quantifying Behavioural Relations

- **Similar to association rules**
 - In treatment → Waiting on results
- **Use similar metrics**
 - Support
 - Confidence
 - Lift

Quantifying Behavioural Relations

- **Expected time**

- Entire process → 14 days
 - In treatment → 10 days
 - Test planned → 4 days
 - Waiting on results → 2 days
 - Results ready → 4 days

- **Support(In treatment) = $\frac{10 \text{ days}}{14 \text{ days}} = 71\%$**



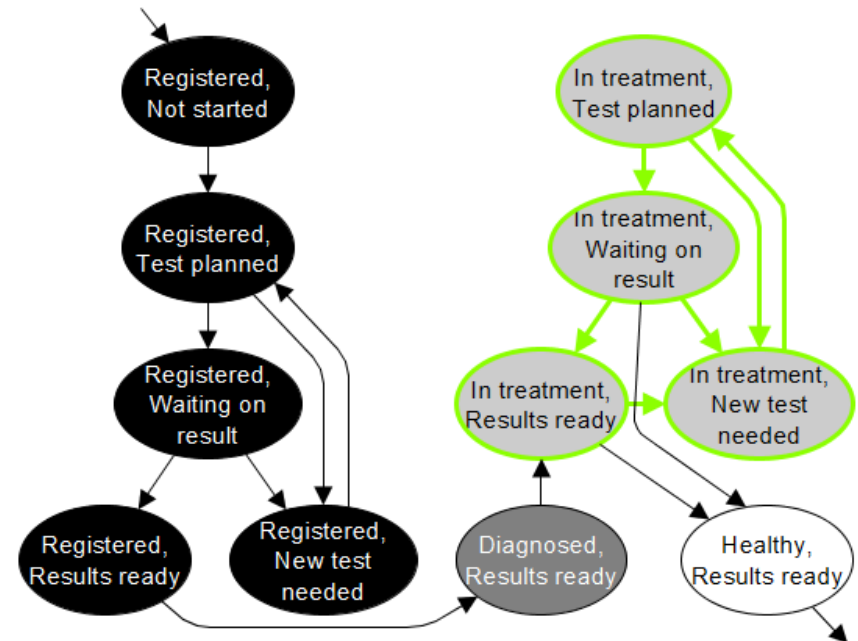
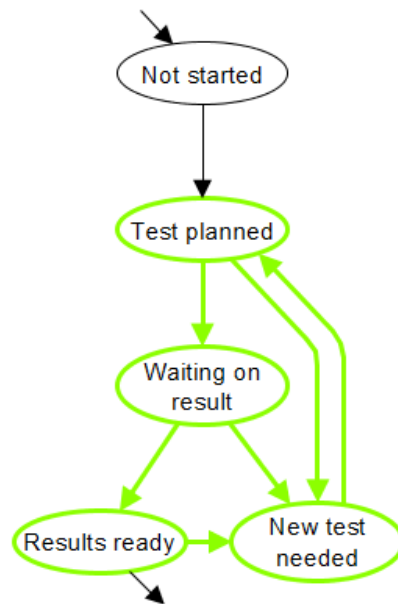
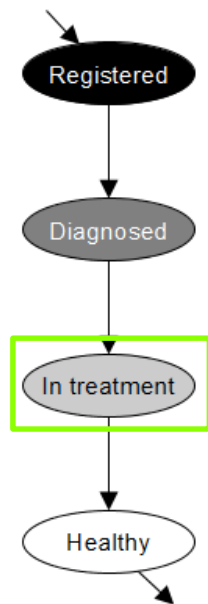
Quantifying Behavioural Relations

- **Expected time**

- **Entire process** → 14 days
 - **In treatment** → 10 days
 - **Test planned** → 4 days
 - **Waiting on results** → 2 days
 - **Results ready** → 4 days

- **Confidence(In treatment → Waiting on results) =**
$$\frac{2 \text{ days}}{10 \text{ days}} = 20\%$$

Quantifying Behavioural Relations





Quantifying Behavioural Relations

- Expected time

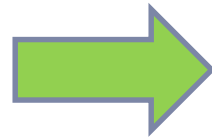
- Entire process → 14 days
 - Waiting on results → 5.5 days (3.5 + 2), $\frac{5,5 \text{ days}}{14 \text{ days}} = 39\%$
 - In treatment → 10 days
 - Test planned → 4 days
 - Waiting on results → 2 days
 - Results ready → 4 days

- Lift(In treatment → Waiting on results) =
 $\frac{20\%}{39\%} = 0,51$

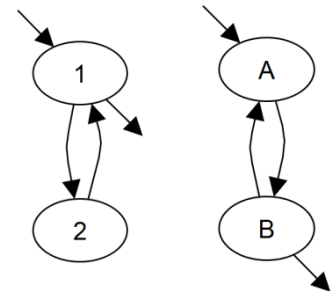
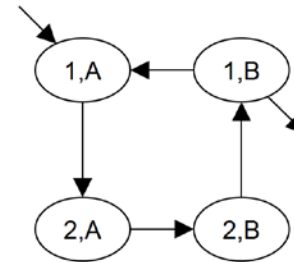
State-based Process Discovery



**State
Log**

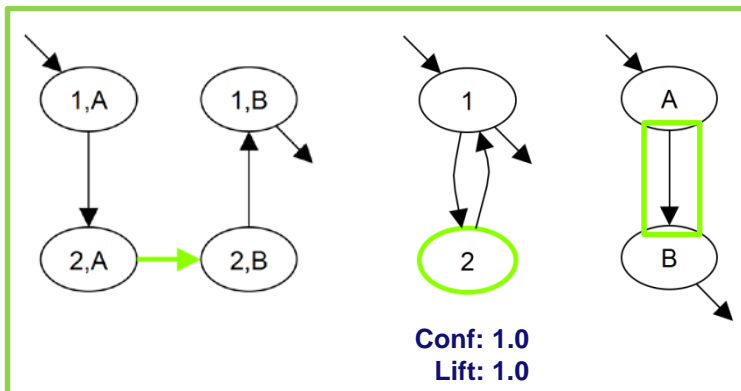
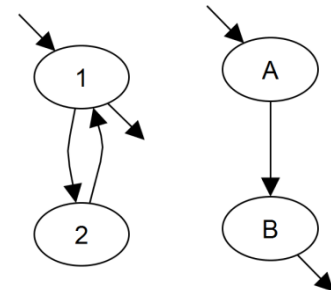
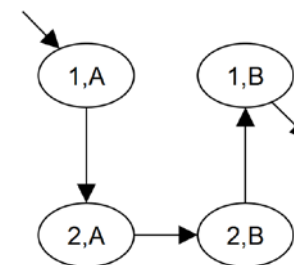
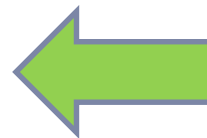


**1. Discovering
Multi-perspective
Models**



**2. Creating
Simplified
Views**

**3. Computing
Behavioural
Relations**

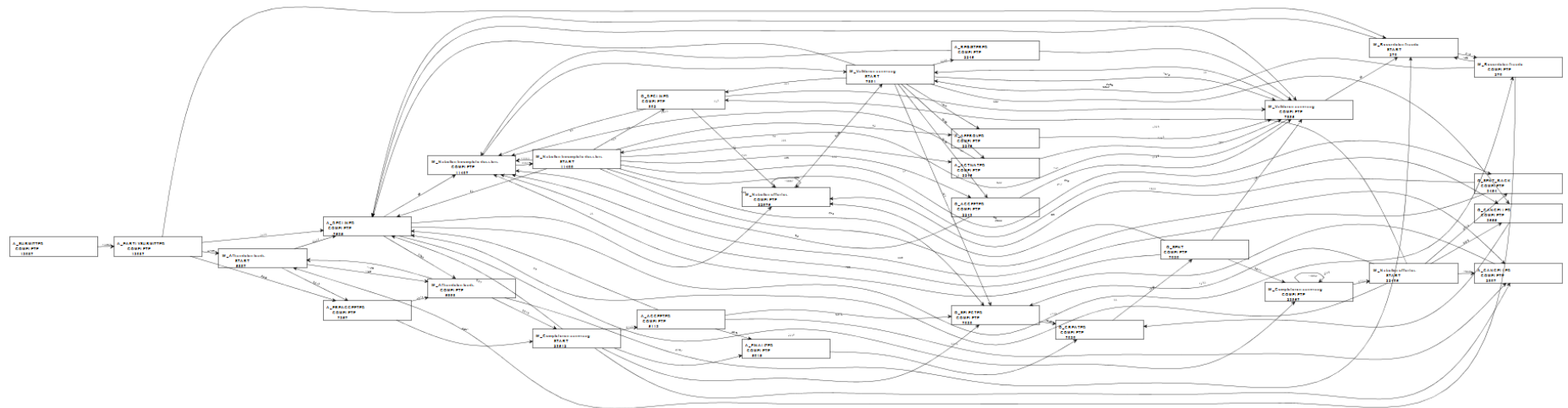




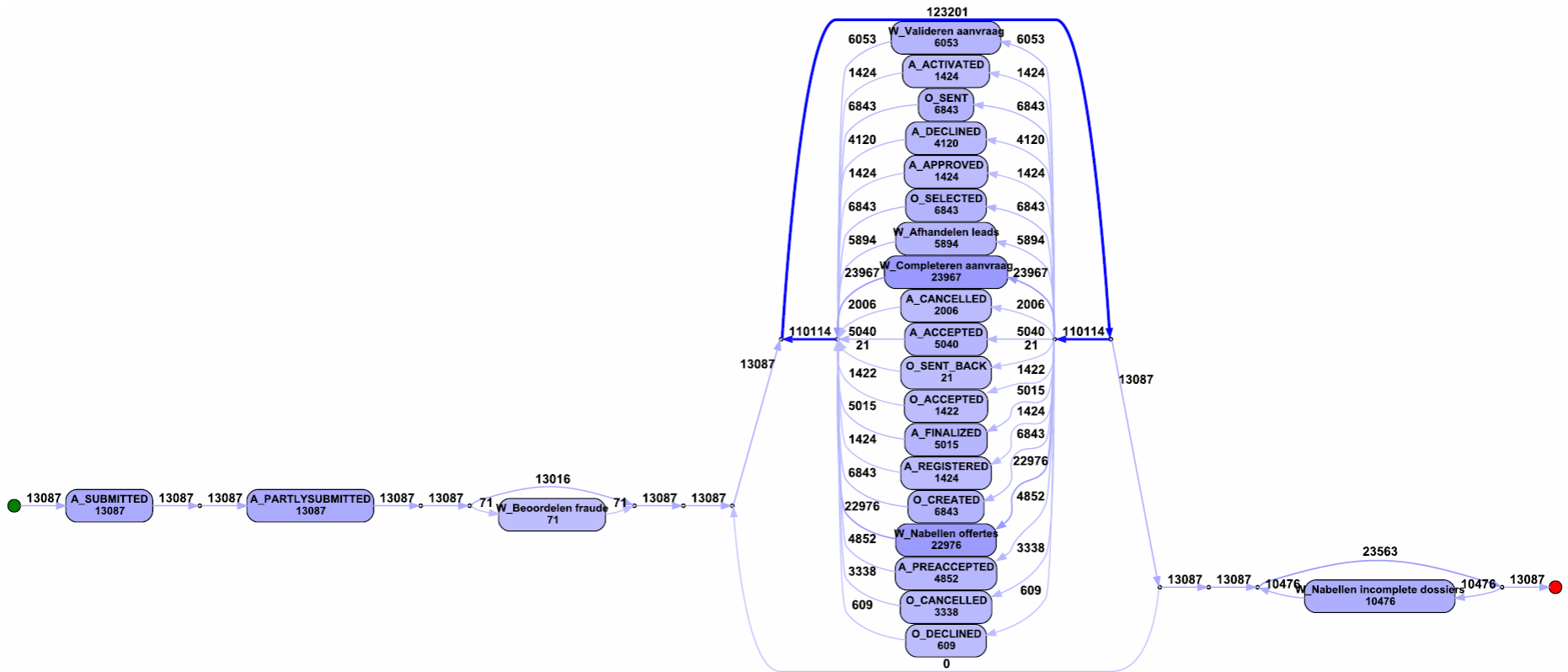
Evaluation

- **BPI Challenge 2012**
 - **Loan application process at a Dutch financial institution**
 - **262.200 events & 13.087 cases**
- **Perspectives:**
 - **State of the client's application – A events**
 - **State of the institution's offer – O events**
 - **What manual work is performed (the state of human resources involved) – W events**

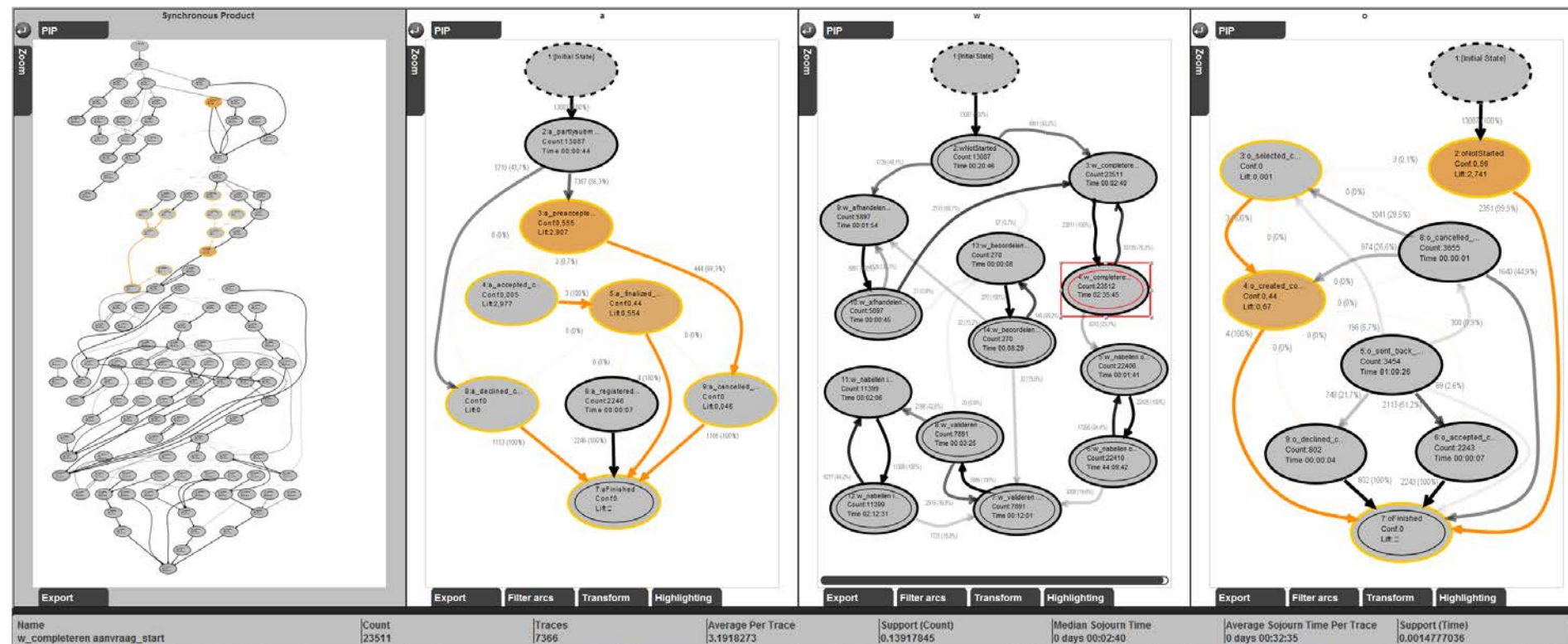
BPI Challenge 2012



BPI Challenge 2012



Evaluation



Conclusions

- **Traditional discovery approaches have difficulties with complex multi-perspective processes**
- **State-based process mining gives simple models**
- **Interactive exploration to investigate the complex relations**
- **Available as a plug-in for ProM 6: CSMMiner**

Future Work

- **Flexible perspective notion**
 - Create sub-logs, e.g. activity clustering, decomposition
 - Allow interactive updates
- **Quantify interestingness**
 - Of relations between perspectives
 - Give suggestions of states and transitions for aggregation or removal
- **Artifact-centric process mining**
 - Perspective with multiple case IDs
 - E.g. 1 order with 3 invoices
- **Show other behavioural relations between perspectives**

Questions?

