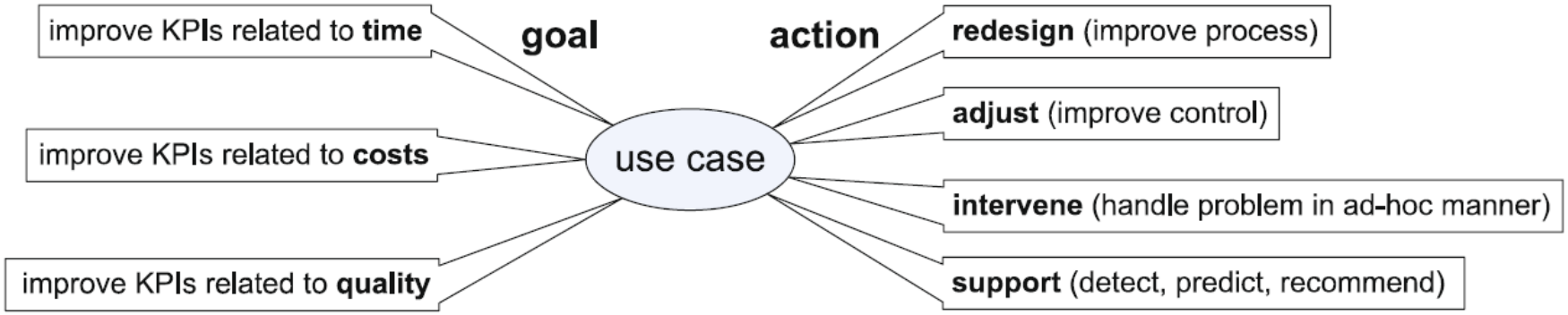
Process Mining

Lesson 9 – Process Mining Topics

Doc. Cognini Riccardo

Analyzing Lasgna Processes

- Lasagna Processes have a clear structure and the flow is not more or less the same
- Input and putput are well defined
- Fitness is more that 80% of cases



Analyzing Spaghetti Processes

- Counterpart of Lasagna Processes: less structured!

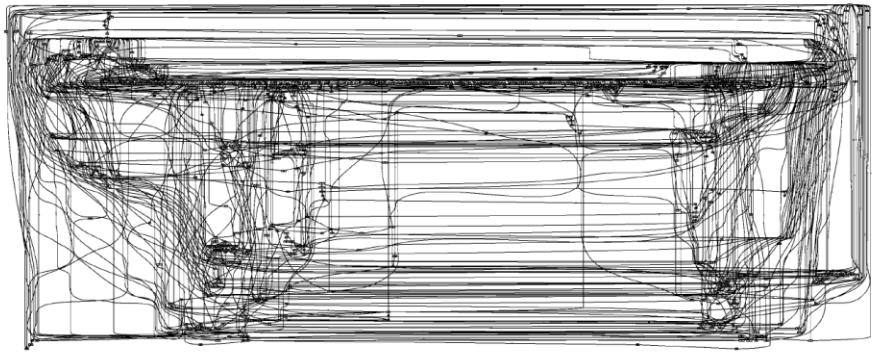
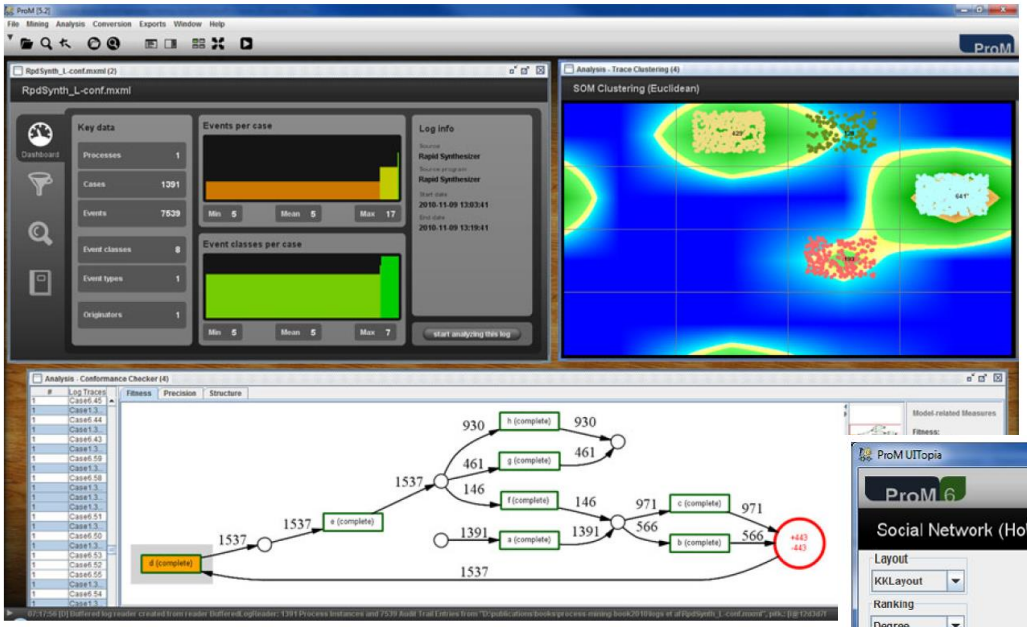
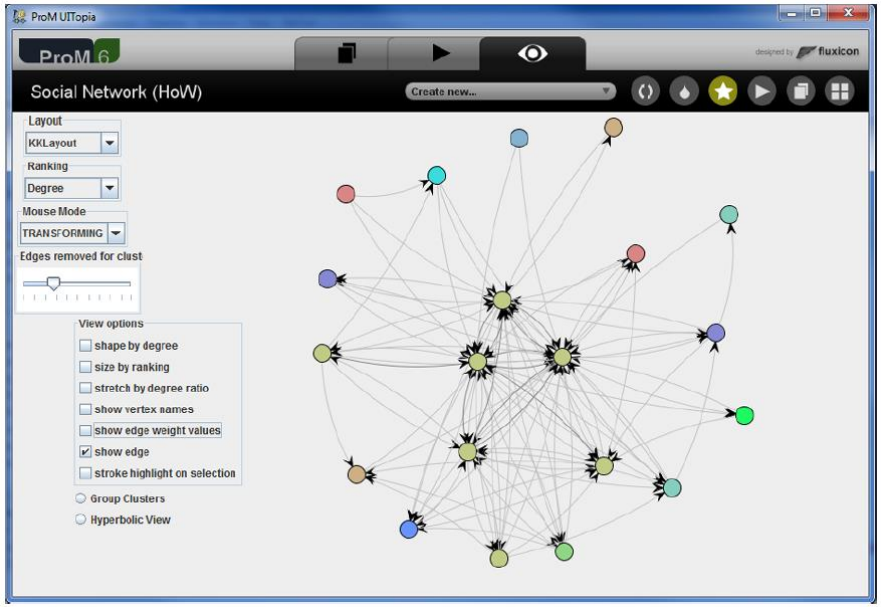


Fig. 12.1 Spaghetti process describing the diagnosis and treatment of 2765 patients in a Dutch hospital. The process model was constructed based on an event log containing 114,592 events. There are 619 different activities (taking event types into account) executed by 266 different individuals (doctors, nurses, etc.)

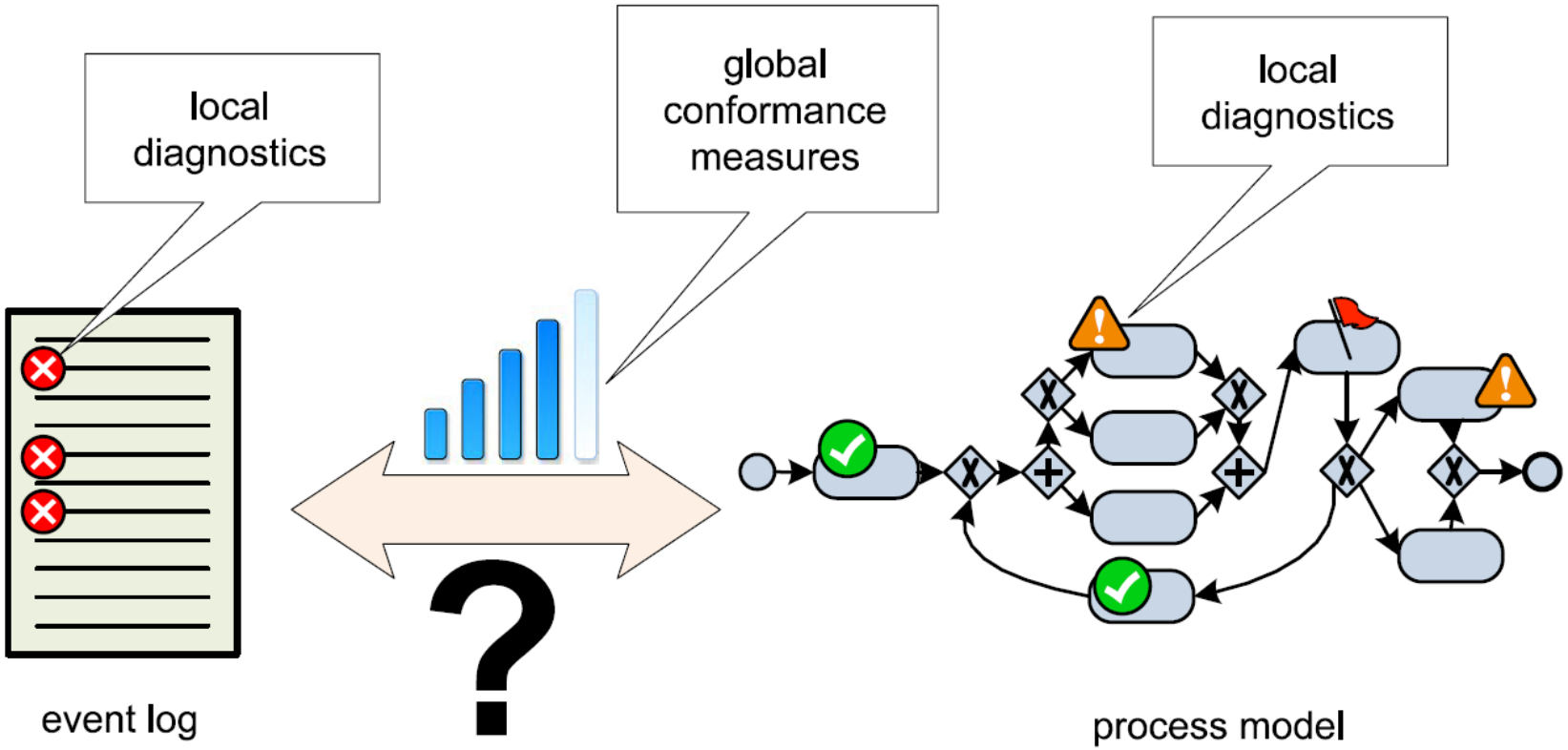
- Filter the activities
- Do not try to do a complete Business Process Model
- Try to mine the more common behaviors (fuzzy mining)

Tools


ProM
RpdSynth_L-conf.mxml
Key data
Processes 1
Cases 1391
Events 7539
Event classes 8
Event types 1
Originators 1
Events per case
Event classes per case
Log info
Rapid Synthesizer
2010-11-09 13:03:41
2010-11-09 13:19:41
start analyzing this log
SOM Clustering (Euclidean)
Analysis - Conformance Checker (4)
Fitness
Precision
Structure
h (complete)
g (complete)
f (complete)
e (complete)
a (complete)
b (complete)
c (complete)
d (complete)
930
461
146
1537
1391
566
971
443


ProM UITopia
ProM 6
fluxicon
Social Network (HoW)
Create new...
Layout
KKLayout
Ranking
Degree
Mouse Mode
TRANSFORMING
Edges removed for clust
View options
shape by degree
size by ranking
stretch by degree ratio
show vertex names
show edge weight values
show edge
stroke highlight on selection
Group Clusters
Hyperbolic View

Conformance Checking



Conformance Checking

- Improve the Business Process model checking if there are wrong behaviors (GOAL)
- Improve the Business Process model including new behaviors that are correct, executed but not modeled (GOAL)
- To check if behaviors are possible Replay technique can be used
- We can check how many traces can be executed in a WF-Net

Conformance Checking

Fitness: «the proportion of behavior in the event log possible according to the model»

We did it in replay game!! But now we should consider also:

- p: produced token
- c: consumed token
- m: mission token
- r: remaining token

Conformance Checking

Fitness of a Trace:

The fitness of a case with trace σ on WF-net N is defined as follows:

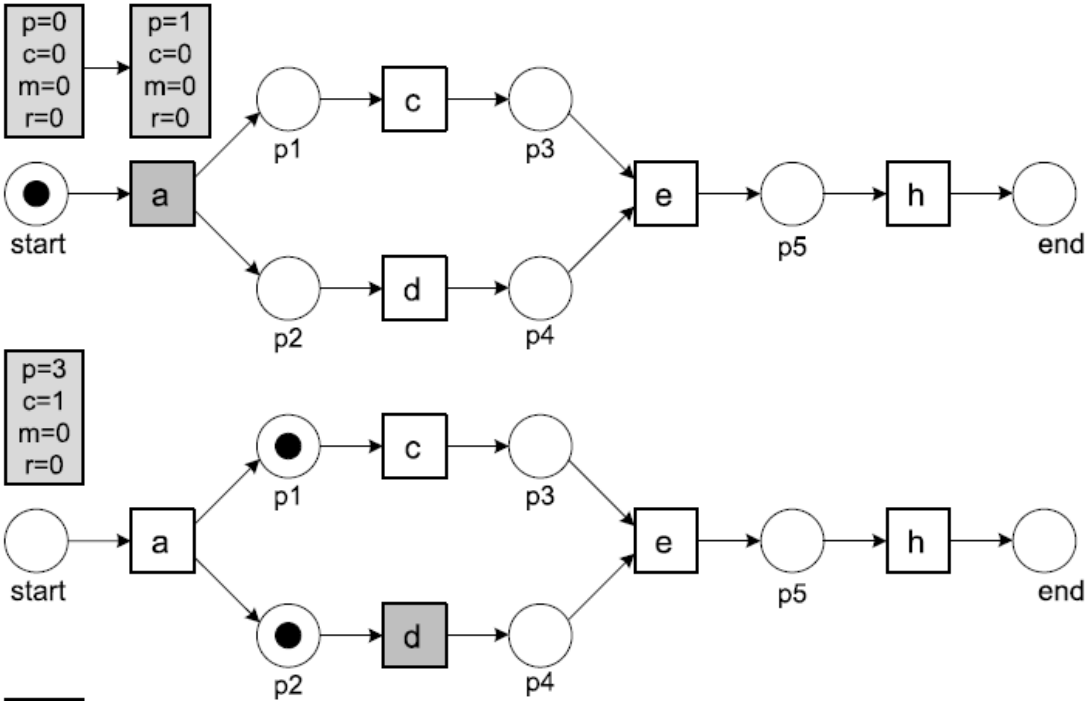
$$fitness(\sigma, N) = \frac{1}{2}\left(1 - \frac{m}{c}\right) + \frac{1}{2}\left(1 - \frac{r}{p}\right)$$

Fitness of a Log:

$$fitness(L, N) = \frac{1}{2}\left(1 - \frac{\sum_{\sigma \in L} L(\sigma) \times m_{N,\sigma}}{\sum_{\sigma \in L} L(\sigma) \times c_{N,\sigma}}\right) + \frac{1}{2}\left(1 - \frac{\sum_{\sigma \in L} L(\sigma) \times r_{N,\sigma}}{\sum_{\sigma \in L} L(\sigma) \times p_{N,\sigma}}\right)$$

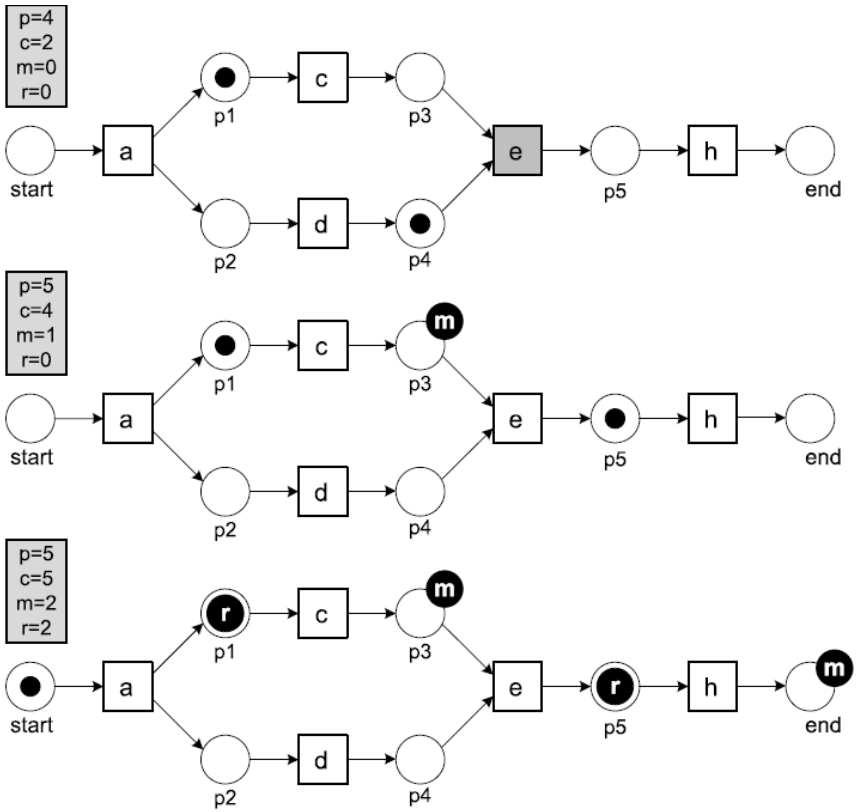
Conformance Checking

$$\sigma_2 = \langle a, b, d, e, g \rangle$$



Conformance Checking

$$\sigma_2 = \langle a, b, d, e, g \rangle$$



QUESTIONS?