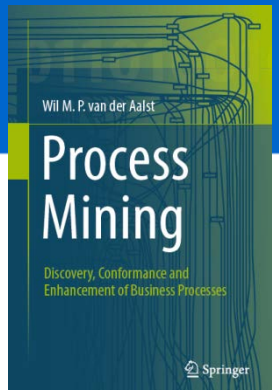
Process Mining: Data Science in Action

Operational Support: Detect, Predict, and Recommend



prof.dr.ir. Wil van der Aalst
www.processmining.org

Technische Universiteit Eindhoven University of Technology

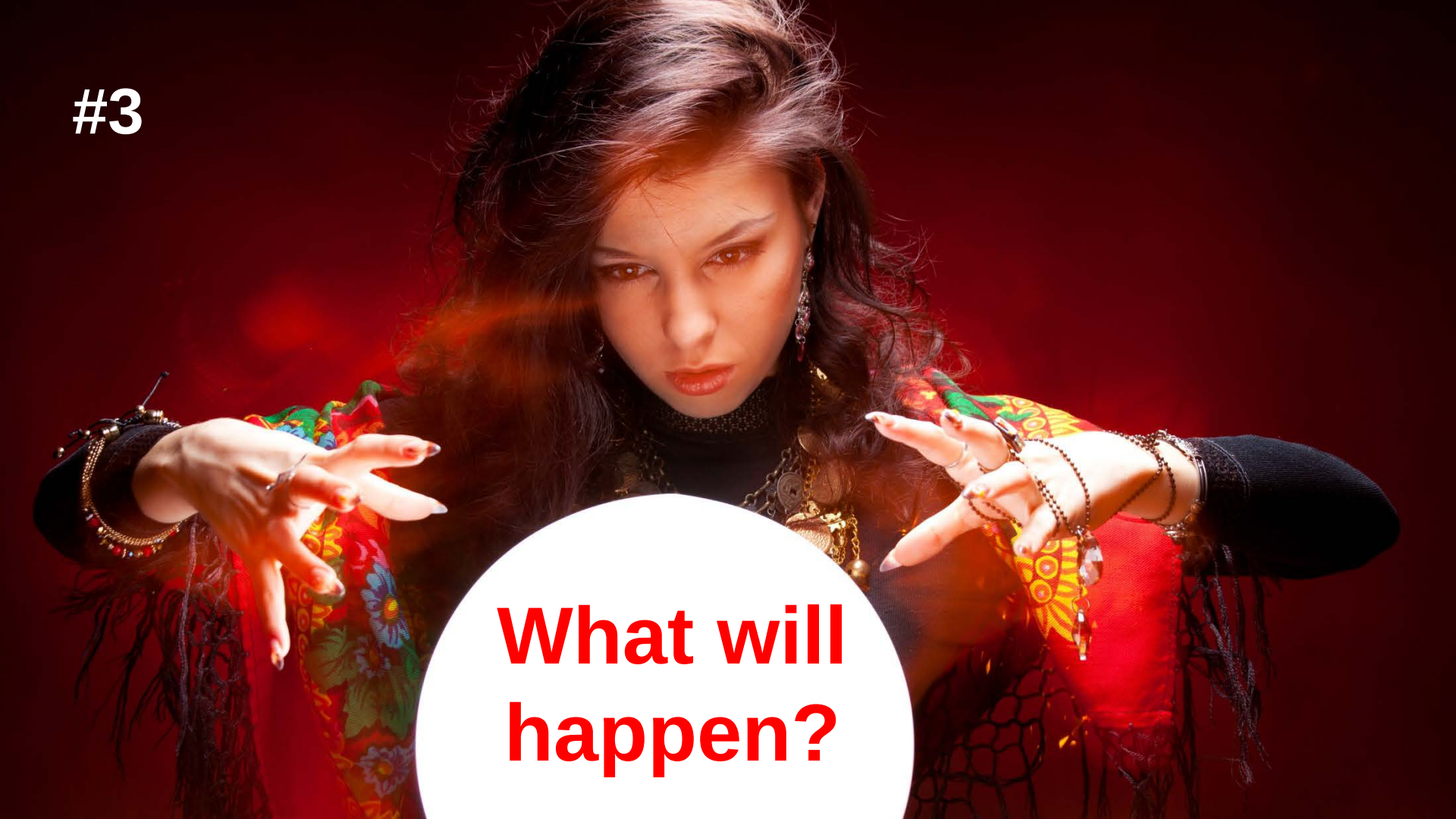
Where innovation starts


Four generic
data science
questions

#1

#2

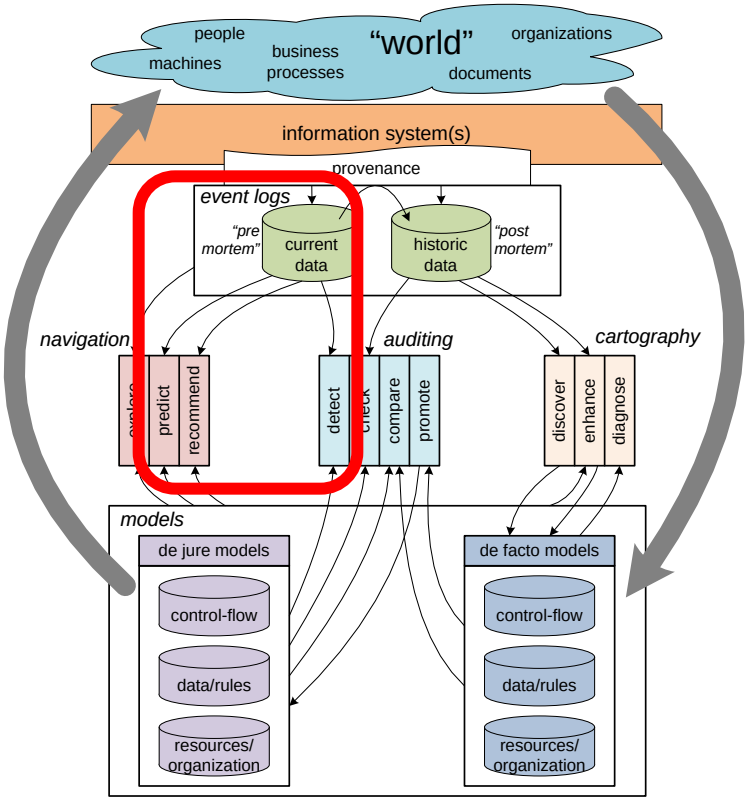
#3



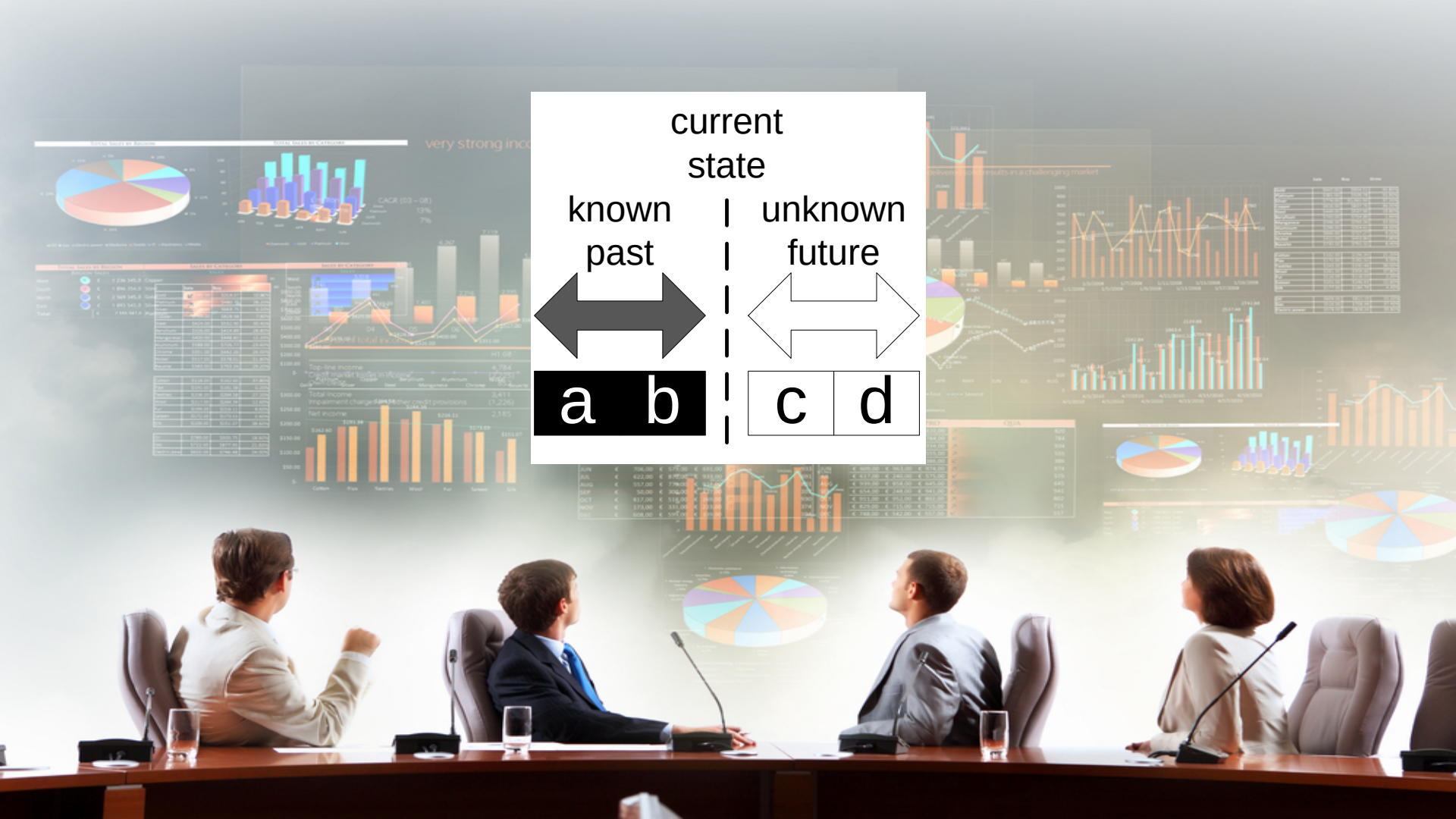
#4

Operational support
Focus on pre mortem data

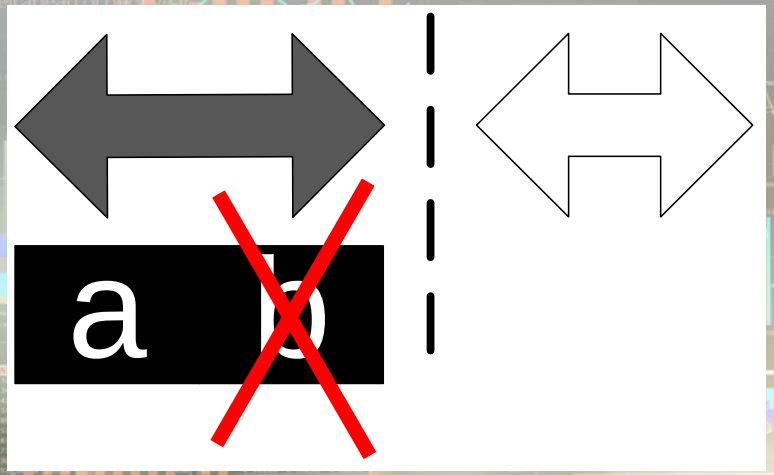
- **Detect**
  - Something is going wrong now!
  - This case is deviating now!
  - The deadline just expired!
- **Predict**
  - When will the case finish?
  - Will the case be rejected?
  - Will the case deviate?
- **Recommend**
  - Which activity should be executed?
  - Who should execute it?



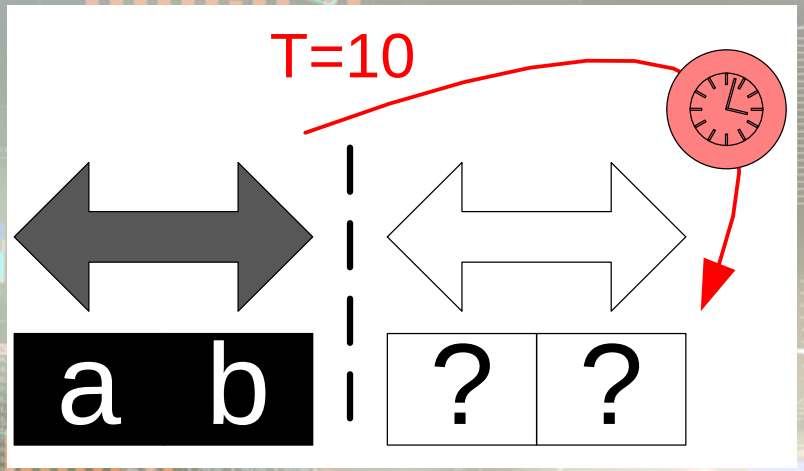
TU/e


current
state
known
past
unknown
future
a
b
c
d

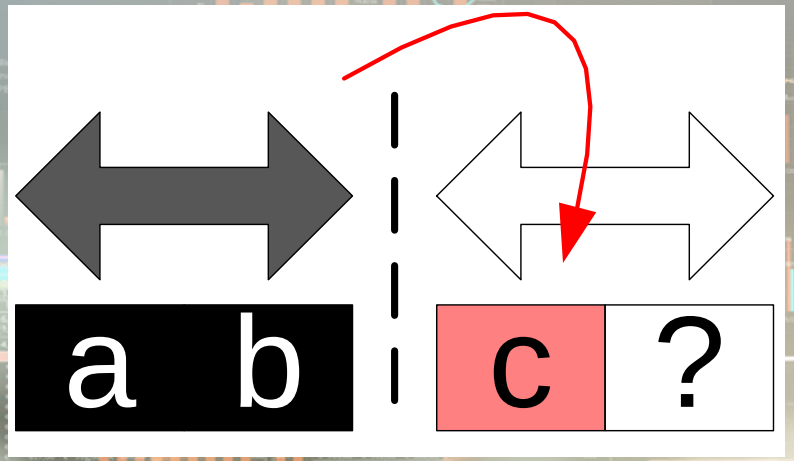
Detect: **b** does not fit the model (not allowed, too late, etc.).



Predict: some prediction is made about the future (e.g. completion date or outcome).

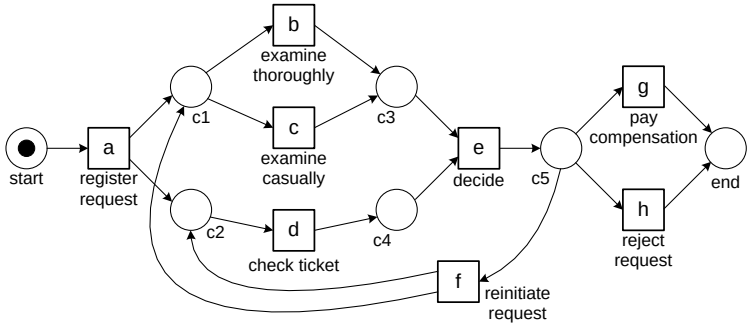


Recommend: based on past experiences **c** is recommended (e.g., to minimize costs).

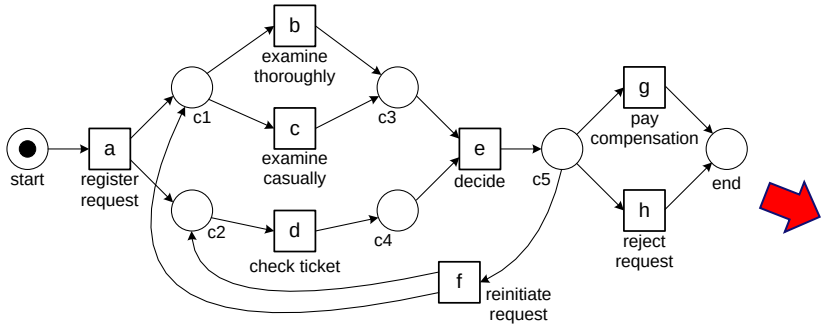


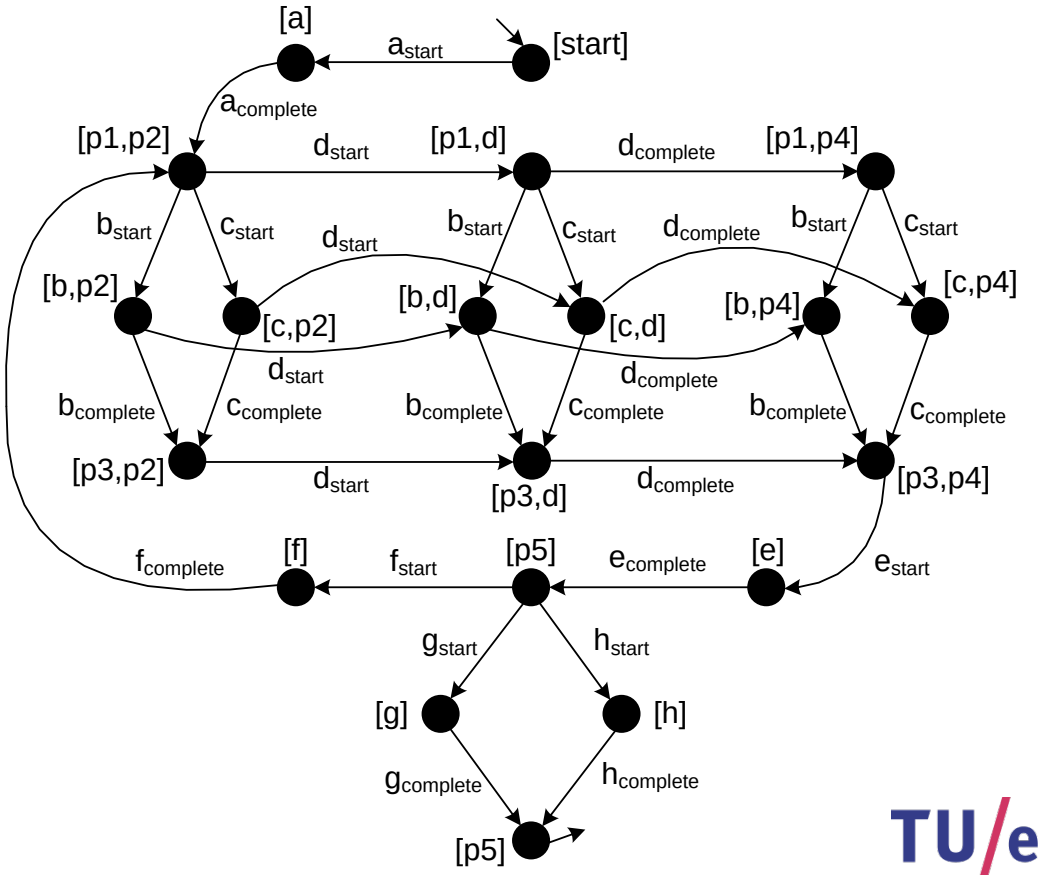
Running example
for simplicity we focus on control-flow and time

case id	trace
1	$\langle a^{12}_{start}, a^{19}_{complete}, b^{25}_{start}, d^{26}_{start}, b^{32}_{complete}, d^{33}_{complete}, e^{35}_{start}, e^{40}_{complete}, h^{50}_{start}, h^{54}_{complete}\rangle$
2	$\langle a^{17}_{start}, a^{23}_{complete}, d^{28}_{start}, c^{30}_{start}, d^{32}_{complete}, c^{38}_{complete}, e^{50}_{start}, e^{59}_{complete}, g^{70}_{start}, g^{73}_{complete}\rangle$
3	$\langle a^{25}_{start}, a^{30}_{complete}, c^{32}_{start}, c^{35}_{complete}, d^{35}_{start}, d^{40}_{complete}, e^{45}_{start}, e^{50}_{complete}, f^{50}_{start}, f^{55}_{complete}, b^{60}_{start}, d^{62}_{start}, b^{65}_{complete}, d^{67}_{complete}, e^{80}_{start}, e^{87}_{complete}, g^{90}_{start}, g^{98}_{complete}\rangle$
...	...

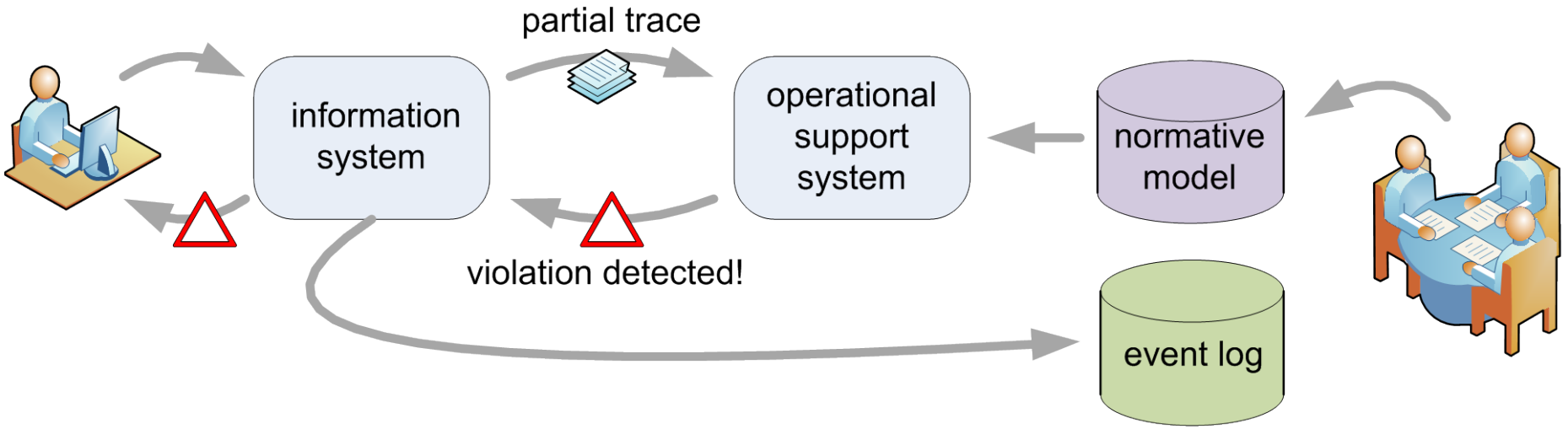


Transition system

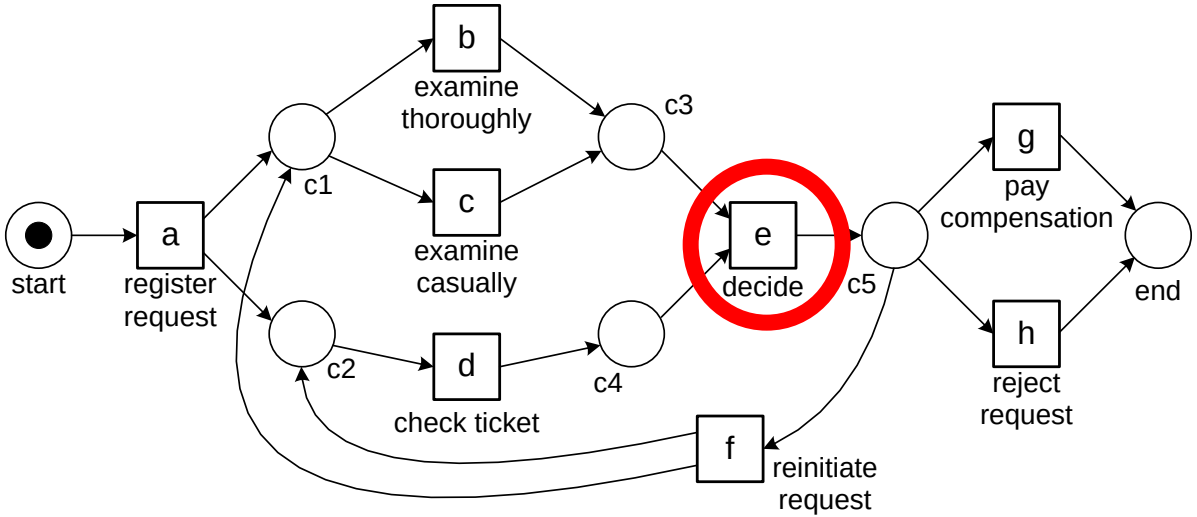




Operational support: Detect



Example

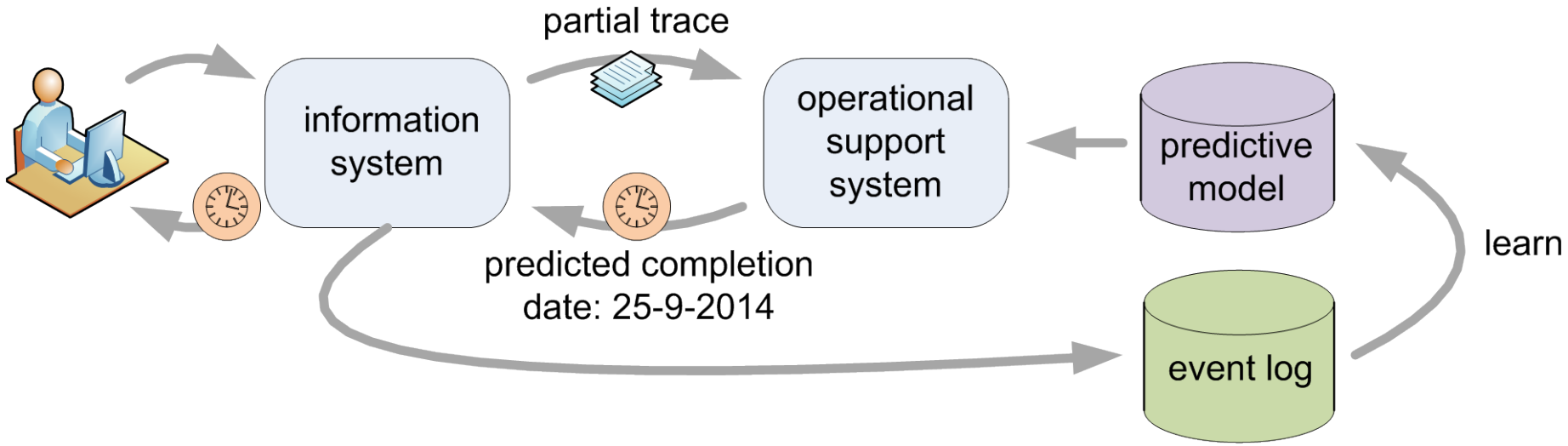


alert @28!!!!

$$\langle a^{12}_{start}, a^{19}_{complete}, b^{25}_{start}, d^{27}_{start}, e^{28}_{start}, \ldots \rangle$$

TU/e

Operational support: Predict



Examples of predictions

The predicted remaining flow time for this case is 14 days.

The predicted probability of meeting the legal deadline is 0.72 for this case.

The predicted probability that person r will work on this case is 0.57.

The predicted probability that activity a will occur is 0.34.

The predicted total cost of this case is €4500.

The predicted probability that this case will be rejected is 0.67.

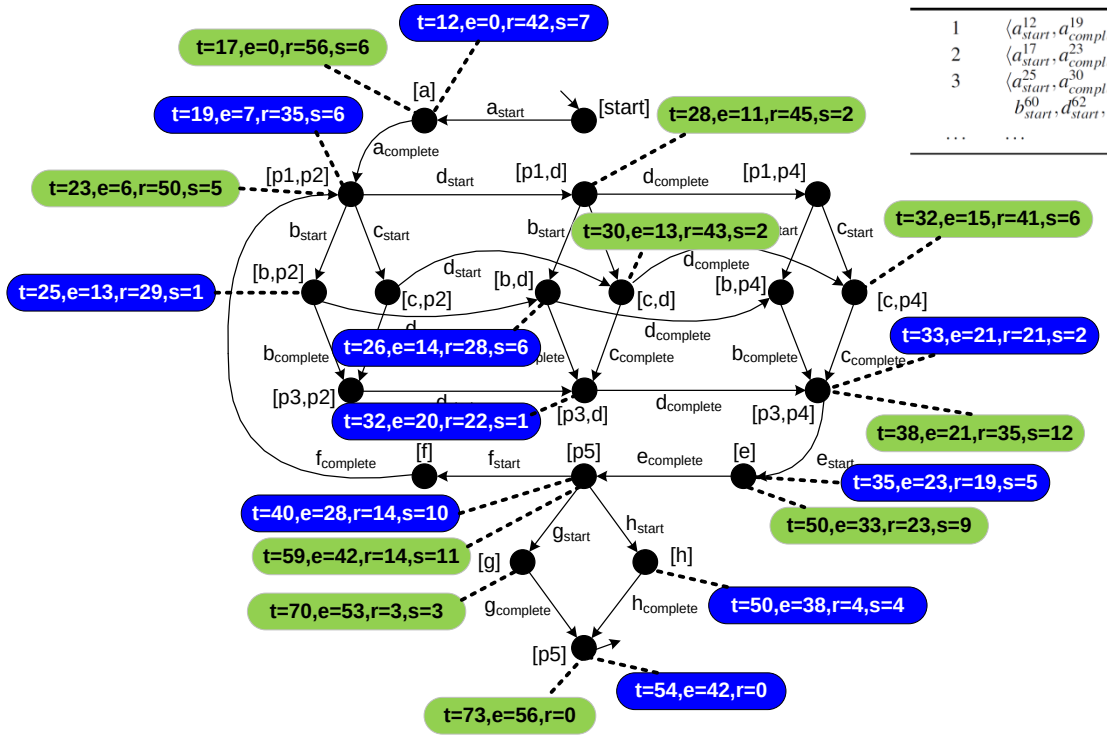
TU/e

Annotated transition system

(based on replay with time, see previous lectures)

case 1

case 2

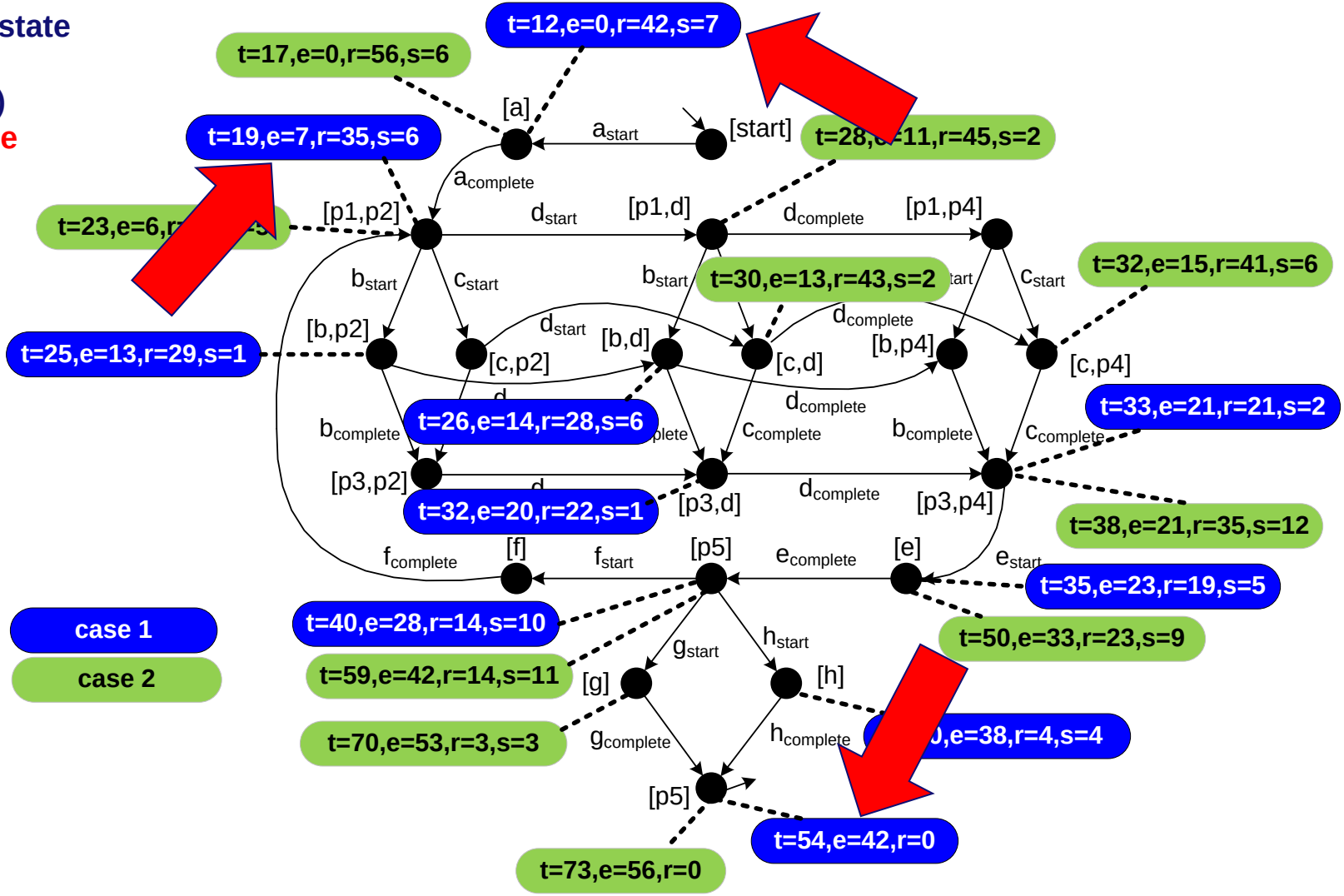


case id	trace
1	$\langle a^{12}_{start}, a^{19}_{complete}, b^{25}_{start}, d^{26}_{start}, b^{32}_{complete}, d^{33}_{complete}, e^{35}_{start}, e^{40}_{complete}, h^{50}_{start}, h^{54}_{complete} \rangle$
2	$\langle a^{17}_{start}, a^{23}_{complete}, d^{28}_{start}, c^{30}_{start}, d^{32}_{complete}, c^{38}_{complete}, e^{50}_{start}, e^{59}_{complete}, g^{70}_{start}, g^{73}_{complete} \rangle$
3	$\langle a^{25}_{start}, a^{30}_{complete}, c^{32}_{start}, c^{35}_{complete}, d^{35}_{start}, d^{40}_{complete}, e^{45}_{start}, e^{50}_{complete}, f^{50}_{start}, f^{55}_{complete}, b^{60}_{start}, d^{62}_{start}, b^{65}_{complete}, d^{67}_{complete}, e^{80}_{start}, e^{87}_{complete}, g^{90}_{start}, g^{98}_{complete} \rangle$
...	...

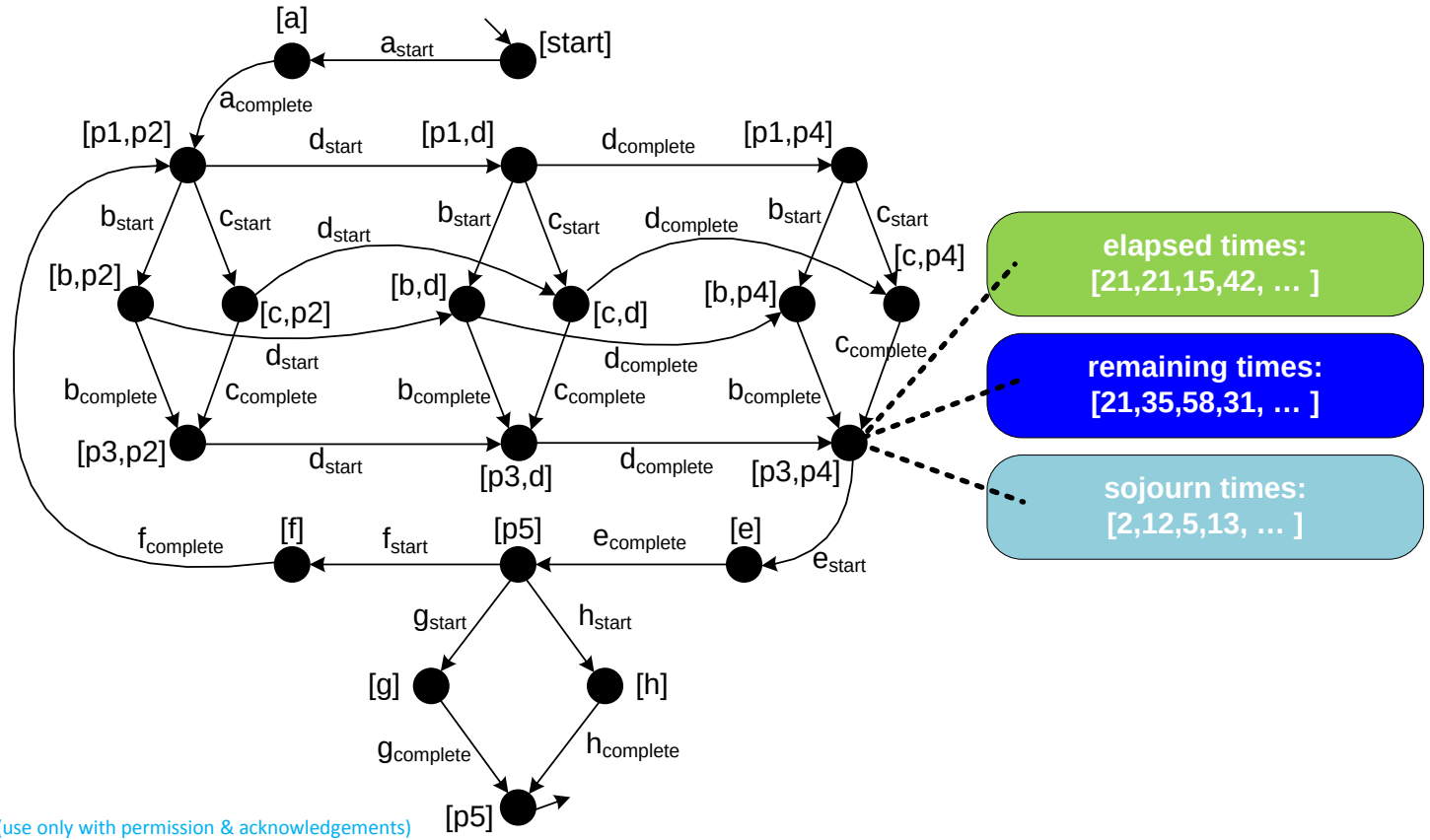
- t = time entering state
- e = elapsed time (since first event)
- r = remaining time (until last event)
- s = sojourn time (time in state)

TU/e

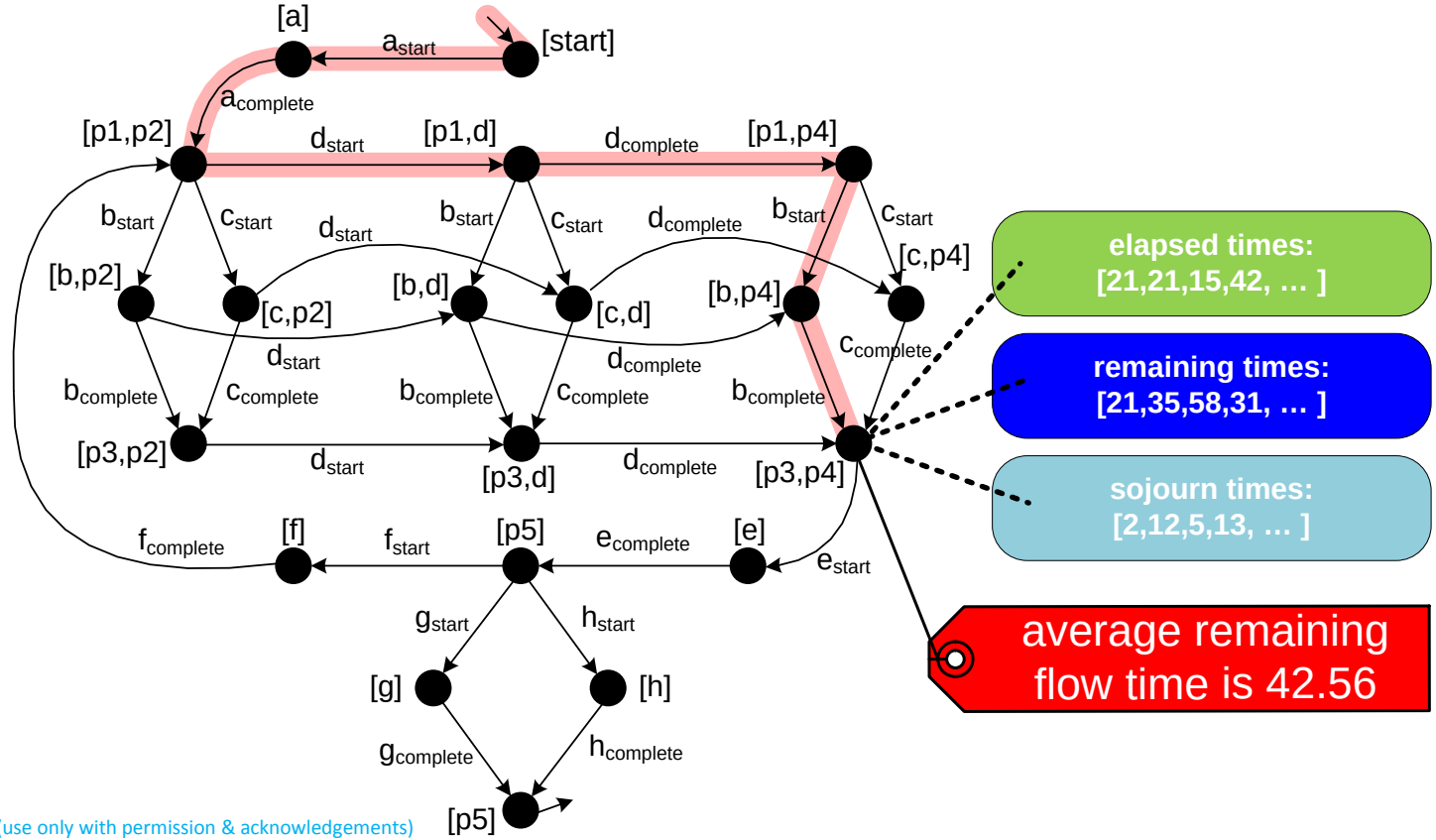
t = time entering state
e = elapsed time
(since first event)
r = remaining time
(until last event)
s = sojourn time
(time in state)



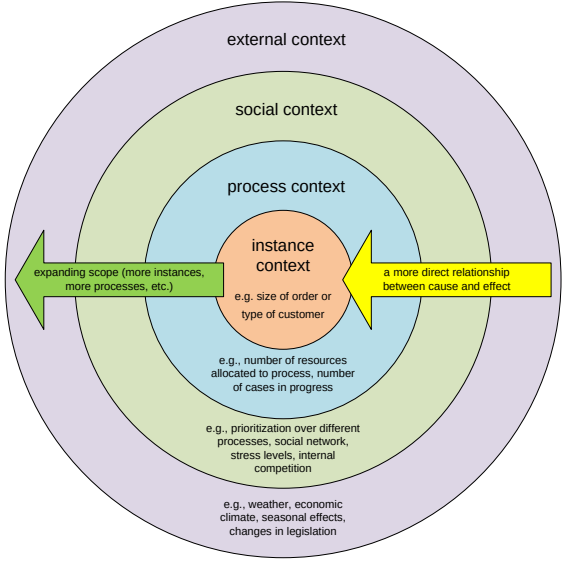
Collect results per state



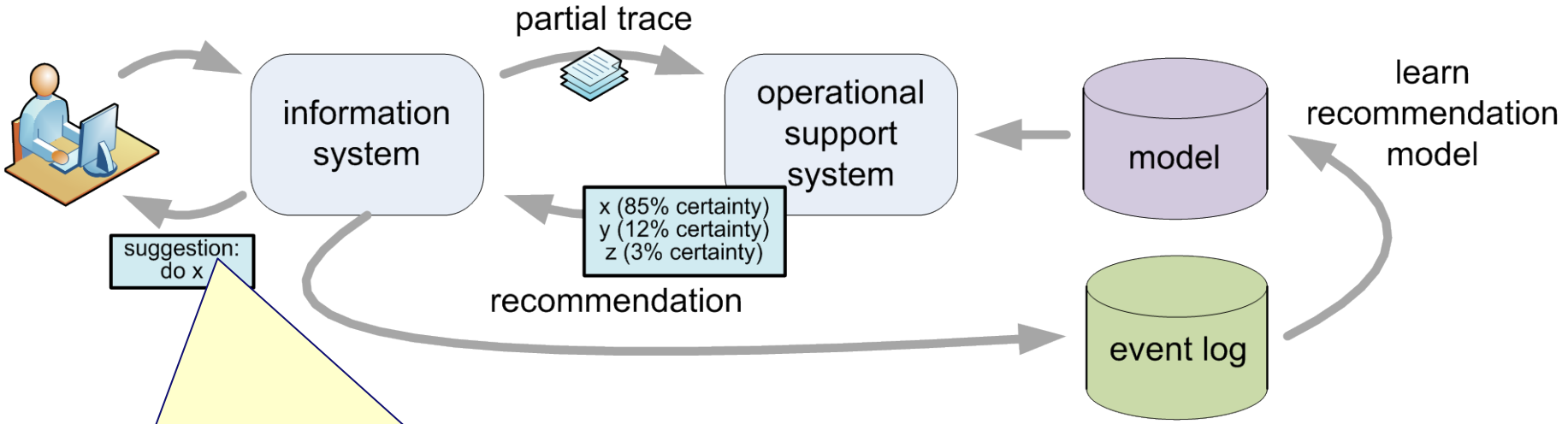
Predict based on current state



Use context information!



Operational support: Recommend



Typical recommendations:

- **Next activity (choice or ordering related).**
- **Suitable resource.**

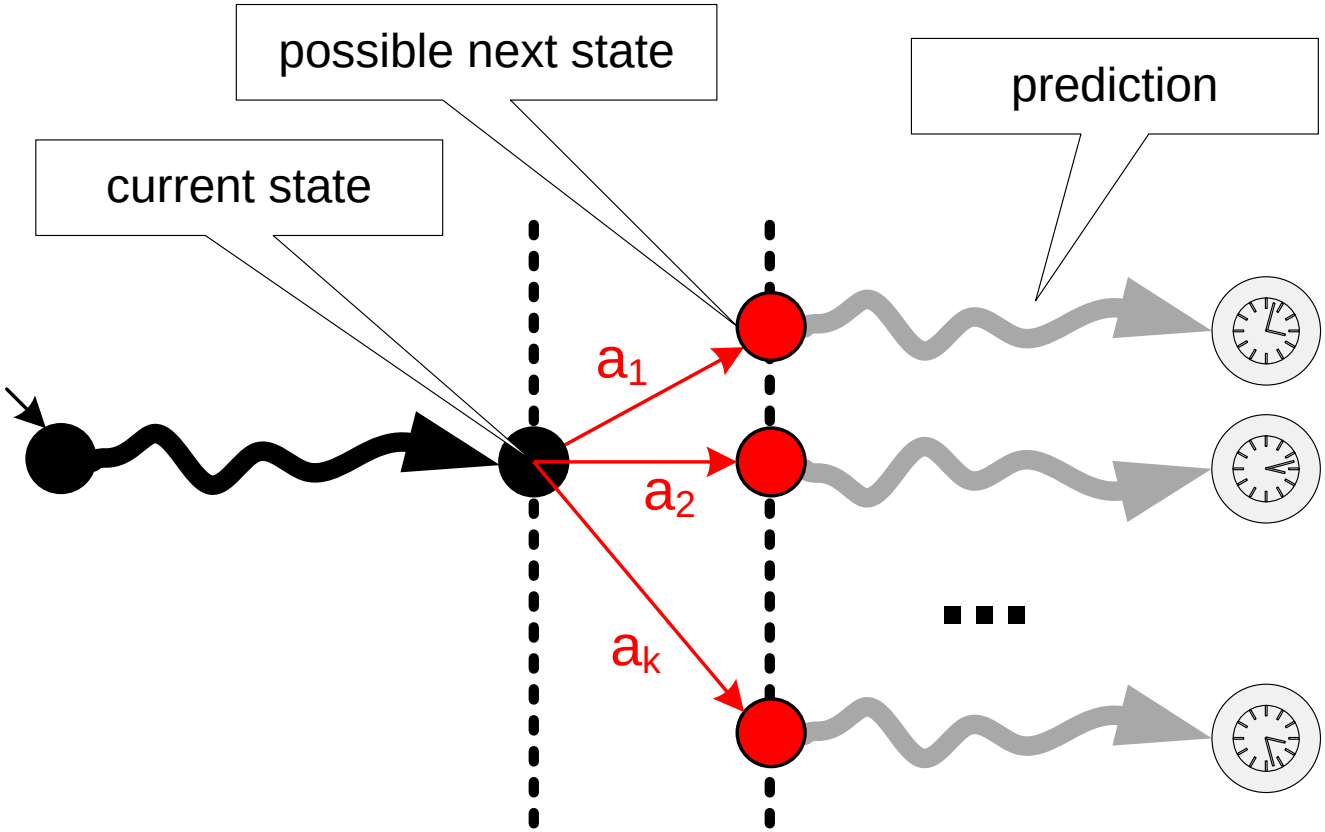
TU/e

A recommendation is always given with respect to a specific goal

- Minimize the remaining flow time.
- Minimize the total costs.
- Maximize the fraction of cases handled within 4 weeks.
- Maximize the fraction of cases that is accepted.
- Minimize resource usage.

TU/e

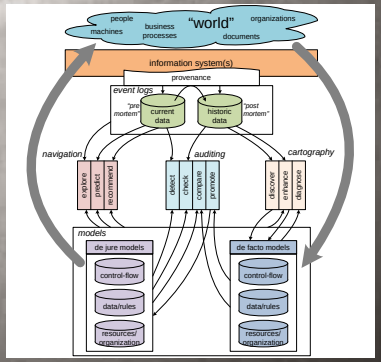
Relation between prediction and recommendation




ability to predict ⇒ ability to recommend

From offline to online: Influence running cases!

Predictive analytics for processes



Part I: Preliminaries

- Chapter 1: Introduction
- Chapter 2: Process Modeling and Analysis
- Chapter 3: Data Mining

Part II: From Event Logs to Process Models

- Chapter 4: Getting the Data
- Chapter 5: Process Discovery: An Introduction
- Chapter 6: Advanced Process Discovery Techniques

Part III: Beyond Process Discovery

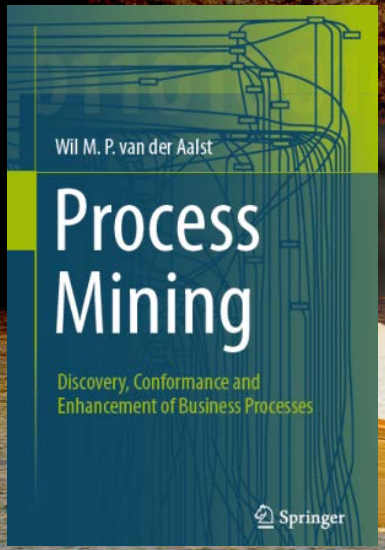
- Chapter 7: Conformance Checking
- Chapter 8: Mining Additional Perspectives
- Chapter 9: Operational Support

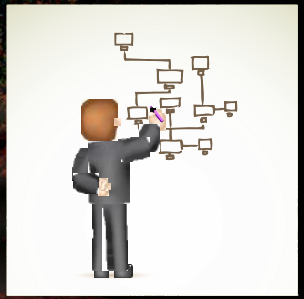
Part IV: Putting Process Mining to Work

- Chapter 10: Tool Support
- Chapter 11: Analyzing "Lasagna Processes"
- Chapter 12: Analyzing "Spaghetti Processes"

Part V: Reflection

- Chapter 13: Cartography and Navigation
- Chapter 14: Epilogue





TU/e