Process mining

ITI8565 Machine Learning
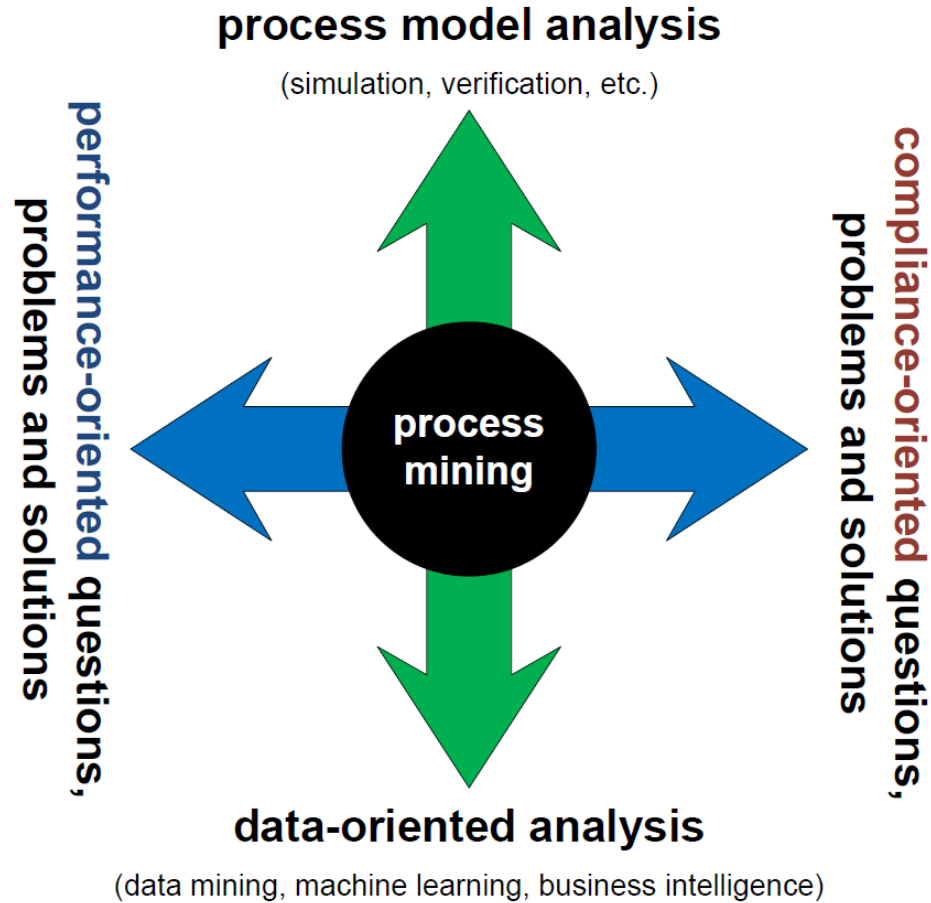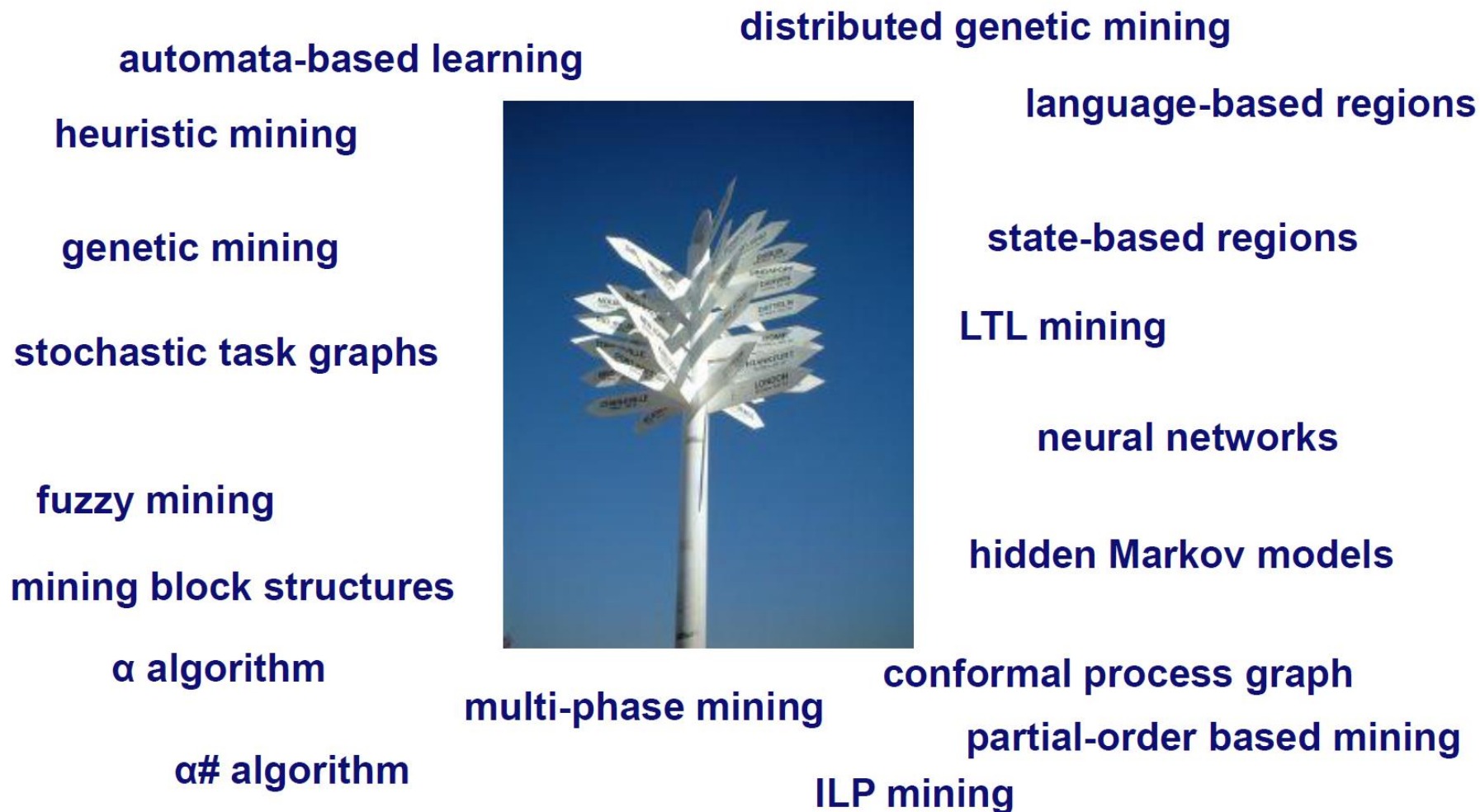
Lecture 15

Juhan Ernits

Background

- Process mining as a field emerged in the naughties and is still developing very fast.
- Builds on work
  - Anil Nerodes's approach to synthesize finite-state machines from example traces (1958)
  - Carl Adam Petri's modelling language that was the first to capture concurrency (1962)
  - Mark Gold's systematic study of different notions of learnability (1967)
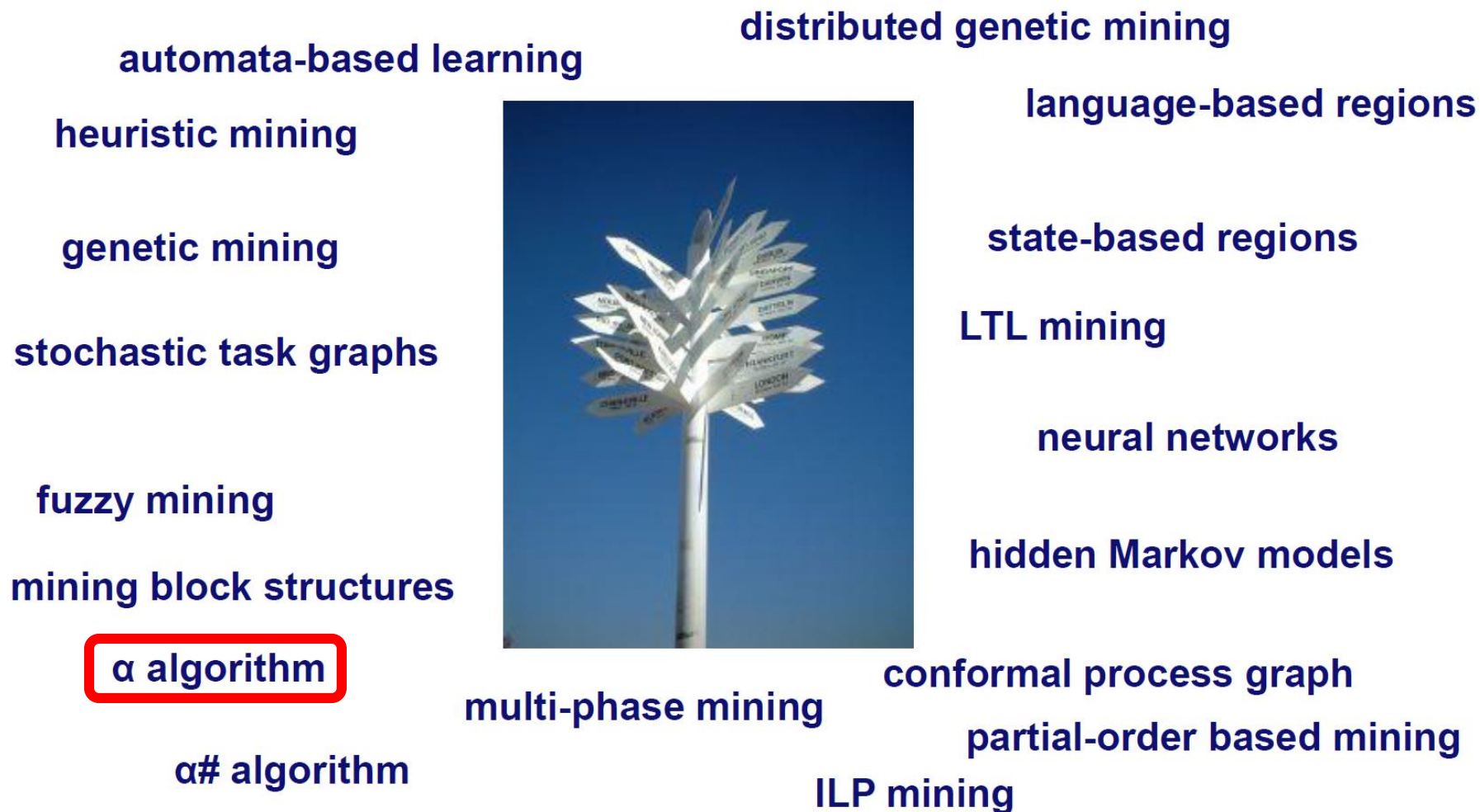
Positioning process mining

process model analysis

(simulation, verification, etc.)

performance-oriented questions, problems and solutions

process mining

compliance-oriented questions, problems and solutions

data-oriented analysis

(data mining, machine learning, business intelligence)

The context of process mining methods

distributed genetic mining

automata-based learning

language-based regions

heuristic mining

genetic mining

state-based regions

stochastic task graphs

LTL mining

neural networks

fuzzy mining

hidden Markov models

mining block structures

α algorithm

conformal process graph

multi-phase mining

partial-order based mining

α# algorithm

ILP mining

The context of process mining methods

automata-based learning

heuristic mining

genetic mining

stochastic task graphs

fuzzy mining

mining block structures

α algorithm

α# algorithm

distributed genetic mining

language-based regions

state-based regions

LTL mining

neural networks

hidden Markov models

multi-phase mining

conformal process graph

partial-order based mining

ILP mining

Typical representational bias

- **(Labeled) Petri Nets, WF-nets, etc.**
- **Subsets of**
  - **BPMN diagrams,**
  - **UML Activity Diagrams,**
  - **Event-Driven Process Chains (EPCs),**
  - **YAWL,**
  - **Statecharts?**
  - **etc.**
- **Transition Systems**
- **(Hidden) Markov Models**
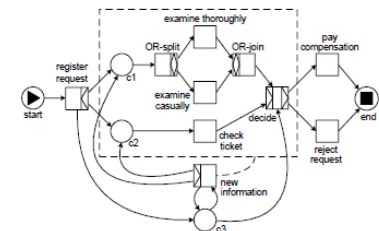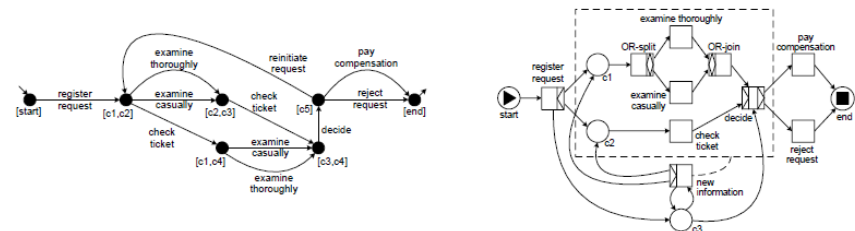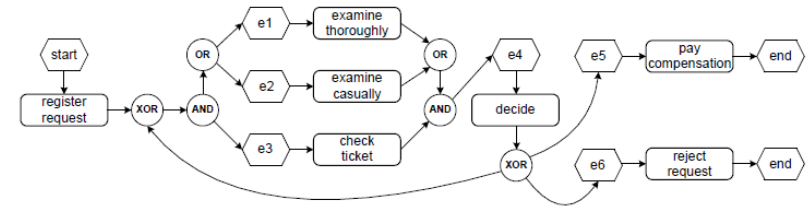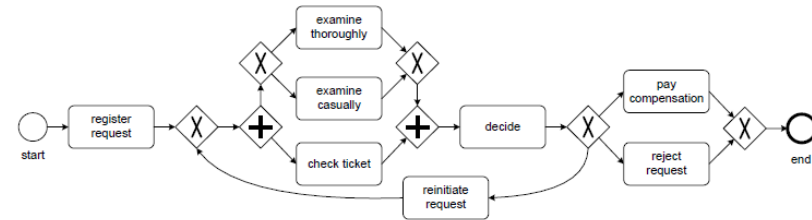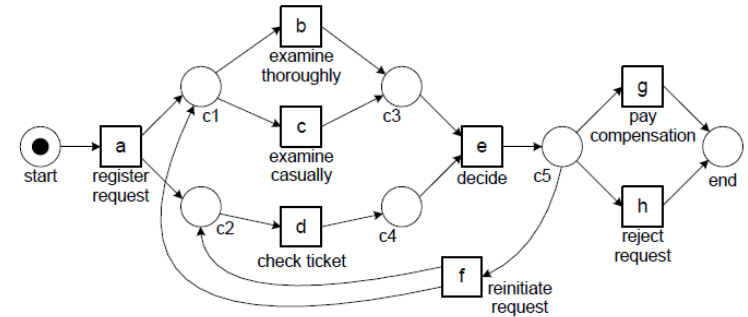
Alternative representational bias

1. **C-nets (XOR/AND/OR-split/join graphs; more likely to be sound due to declarative semantics).**

2. **Declare models (constraint based, grounded in LTL; anything is possible unless forbidden)**

3. **Process Trees (similar to subsets of various process algebras; sound by structure)**
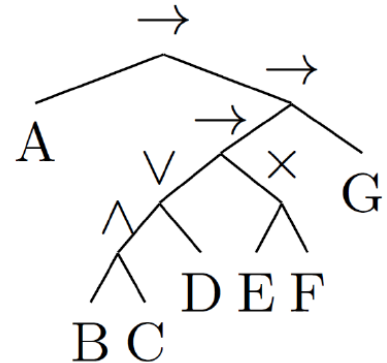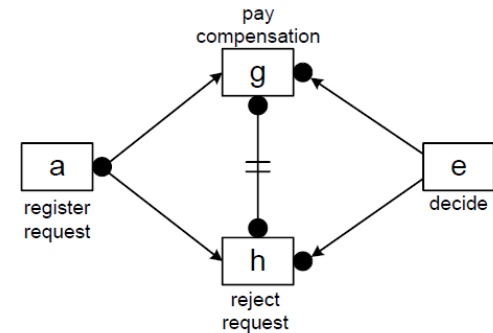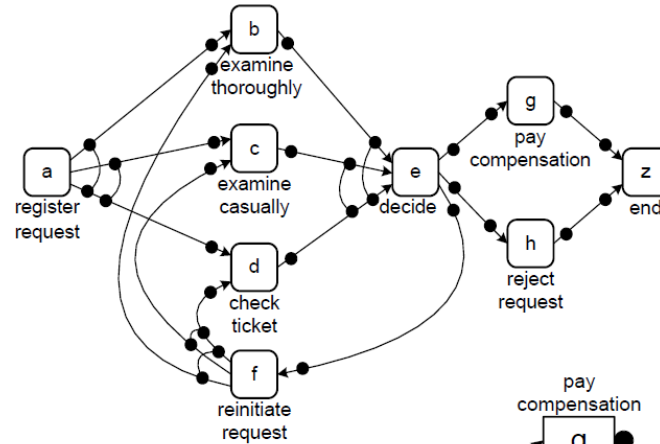
Tool support

- Prom is available from www.processmining.org
- Various plugins made available within the framework.
- Standard log format (XES). Tools to convert logs to that format available from above.

The alpha algorithm

- The rest of the lecture is based on slides for Prom book Chapter 5: http://www.processmining.org/_media/processminingbook/process_mining_chapter_05_process_discovery.pdf

Acknowledgements

- Some slides used in this presentation were taken from http://www.processmining.org/_media/presentations/2013/4-siks-lets-play-2013-discovery-wvda.pdf