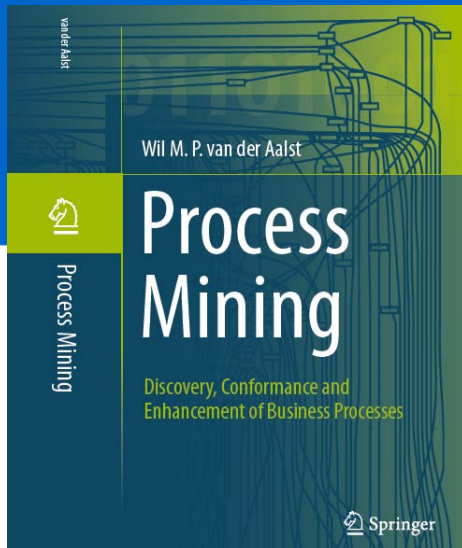


# Chapter 2

## Process Modeling and Analysis

prof.dr.ir. Wil van der Aalst  
[www.processmining.org](http://www.processmining.org)



**TU/e** Technische Universiteit  
Eindhoven  
University of Technology

Where innovation starts

# Overview

Chapter 1  
Introduction

---

## *Part I: Preliminaries*

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”

---

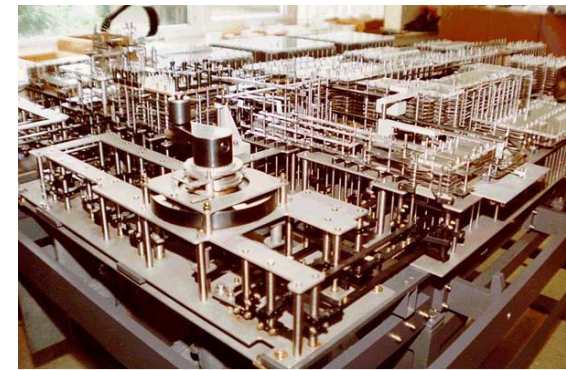
## *Part V: Reflection*

Chapter 13  
Cartography and  
Navigation

Chapter 14  
Epilogue

# Productivity improvements

- **Adam Smith** (1723-1790) showed the advantages of the division of labor.
- **Frederick Taylor** (1856-1915) introduced the initial principles of scientific management.
- **Henry Ford** (1863-1947) introduced the production line for the mass production of “black T-Fords”.
- Since 1950 **computers and digital communication infrastructures** are the most dominant factor influencing business processes and their management.



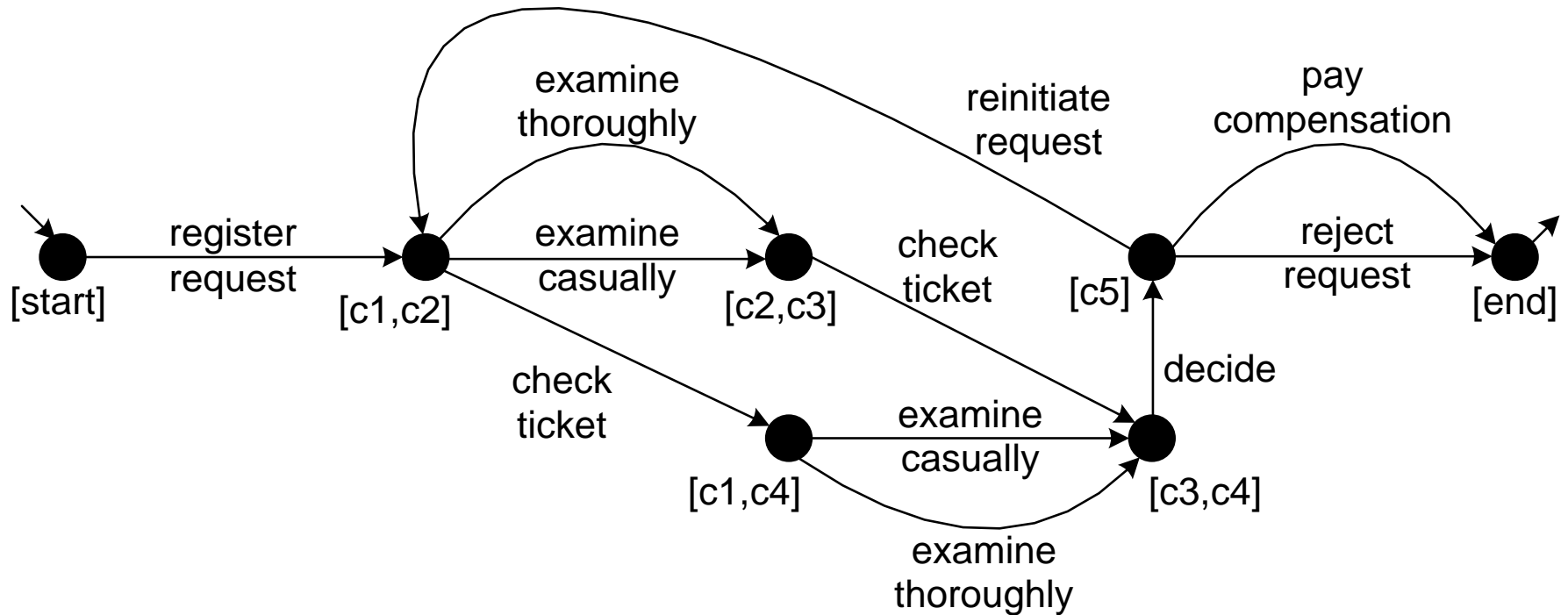
# Role of models

- **Operations management, and in particular operation research, is a branch of management science heavily relying on modeling.**
- **Models are used to reason *about processes* (redesign) and to make decisions *inside processes* (planning and control).**
- **A process model may be used to discuss responsibilities, analyze compliance, predict performance using simulation, and configure a WFM system.**

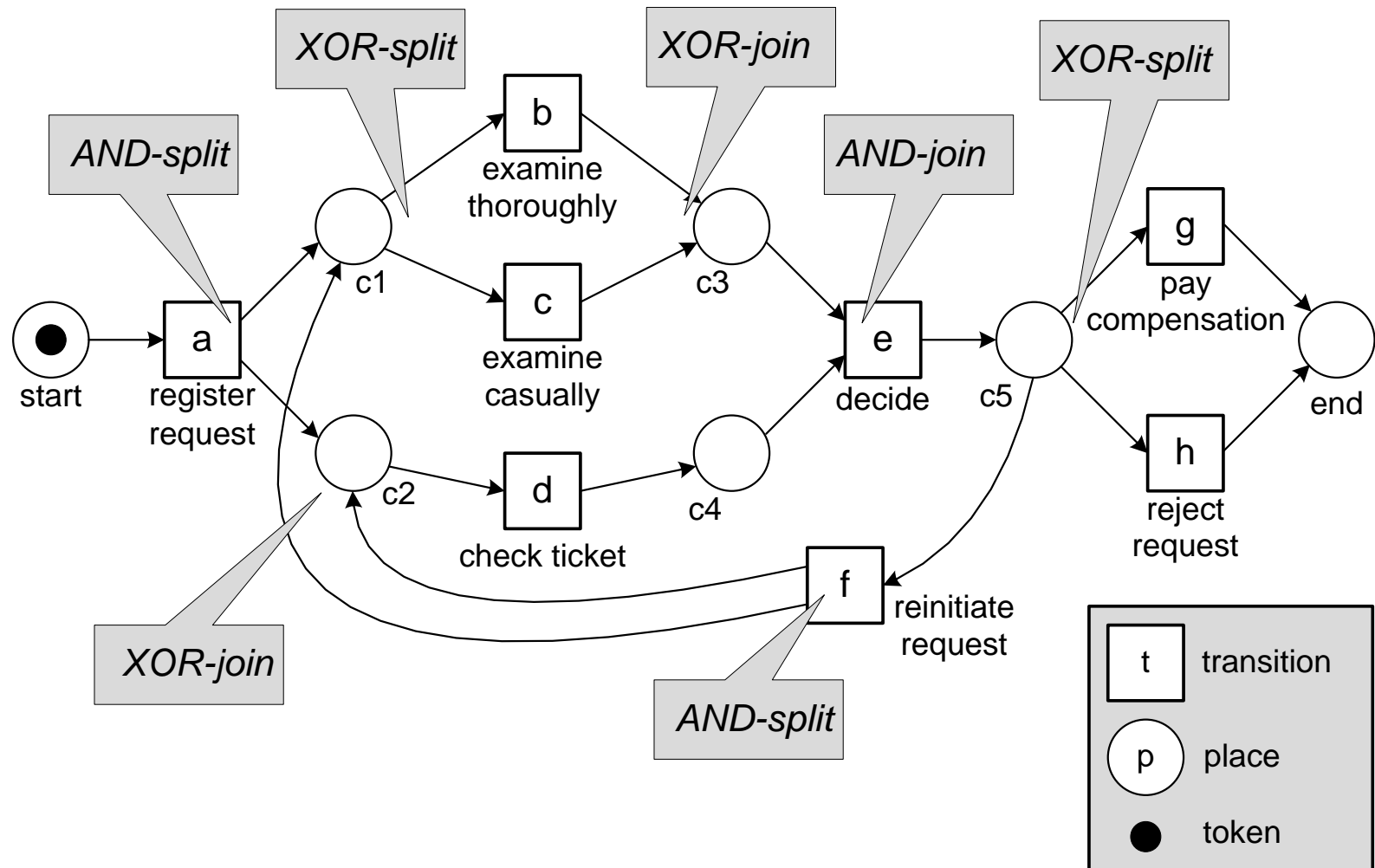
# Problems of models

- **The model describes an idealized version of reality.**
- **Inability to adequately capture human behavior.**
- **The model is at the wrong abstraction level.**
- **Therefore, we advocate the use of event data:**
  - **Process mining allows for the extraction of models based on facts.**
  - **Moreover, process mining does not aim at creating a single model of the process.**
  - **Instead, it provides various views on the same reality at different abstraction levels.**
  - **For example, users can decide to look at the most frequent behavior to get a simple model (“80% model”).**
  - **However, they can also inspect the full behavior by deriving the “100% model” covering all cases observed.**

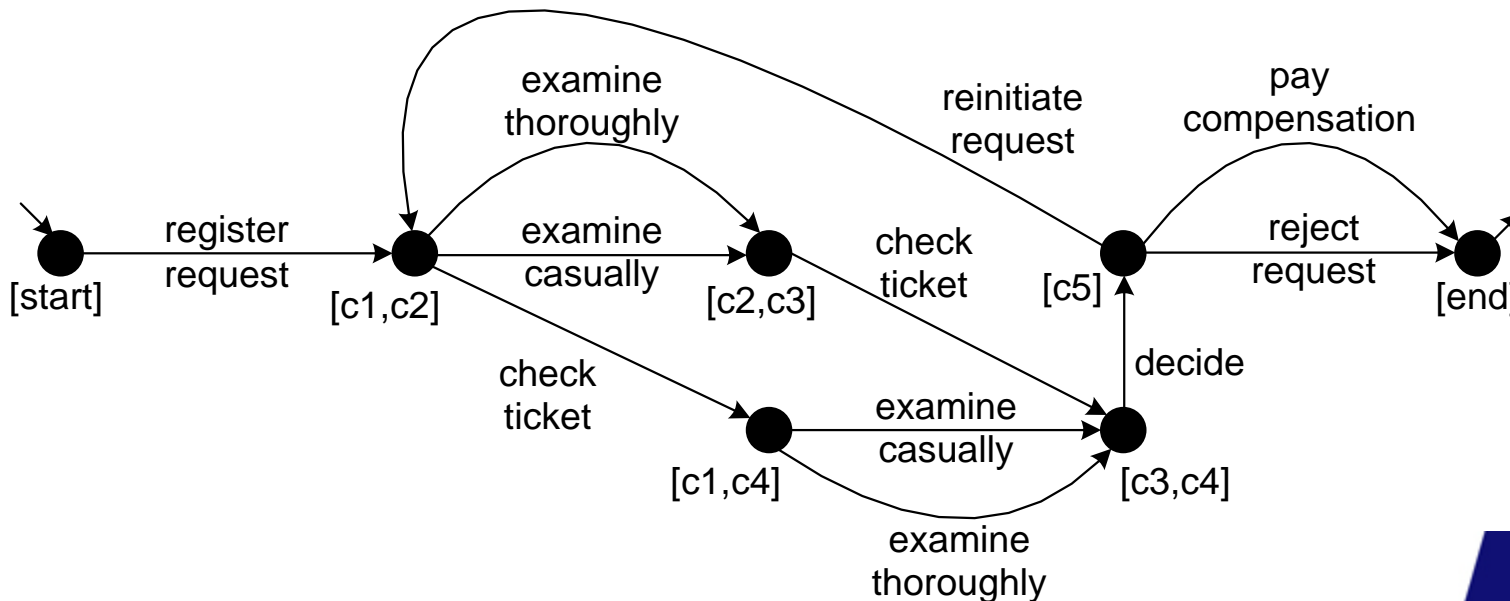
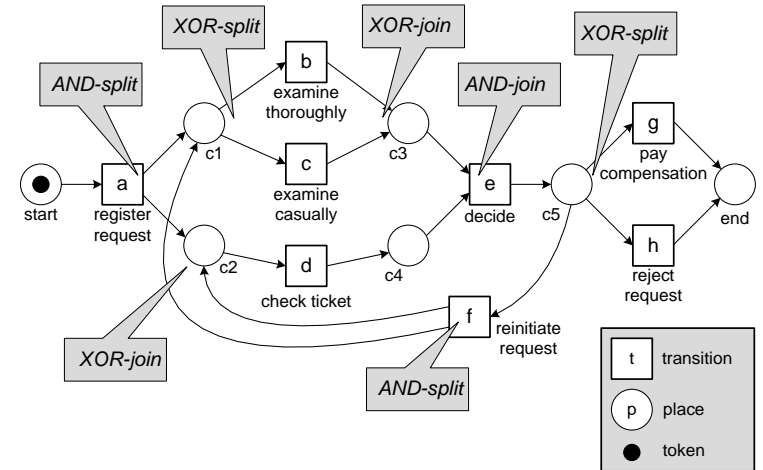
# Transition systems



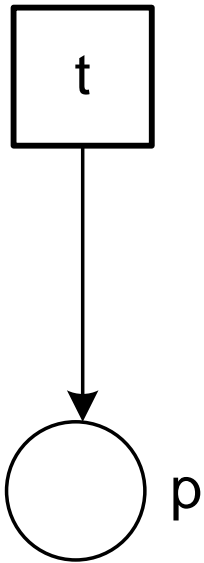
# Petri nets



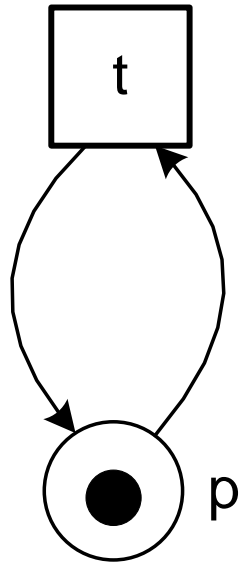
# Reachability graph



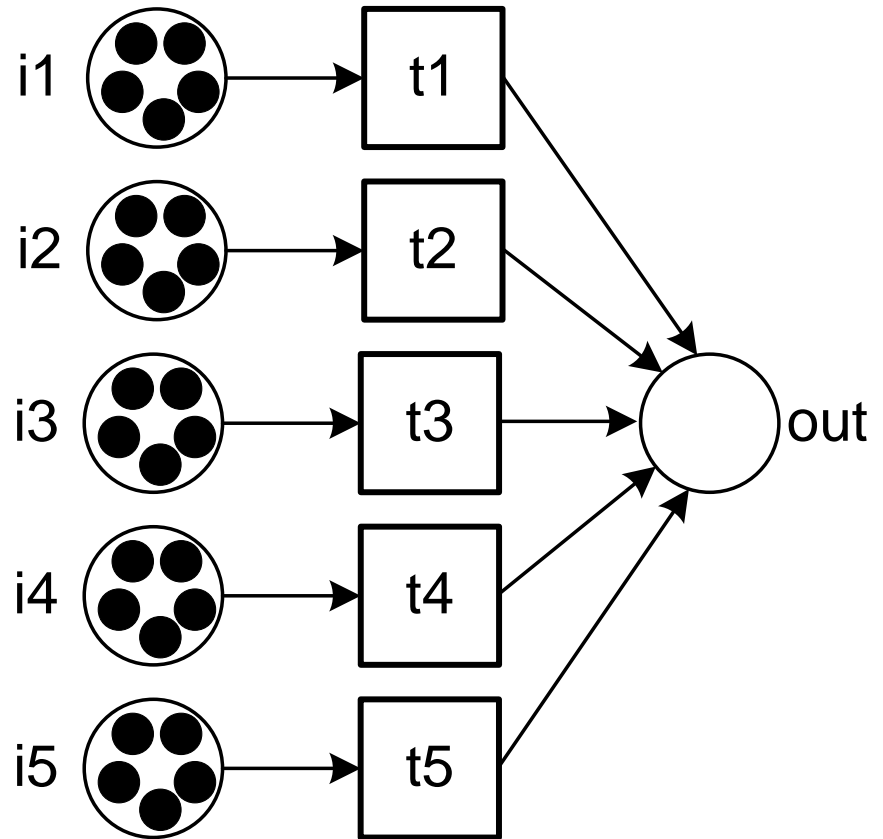
# How many states?



(a)

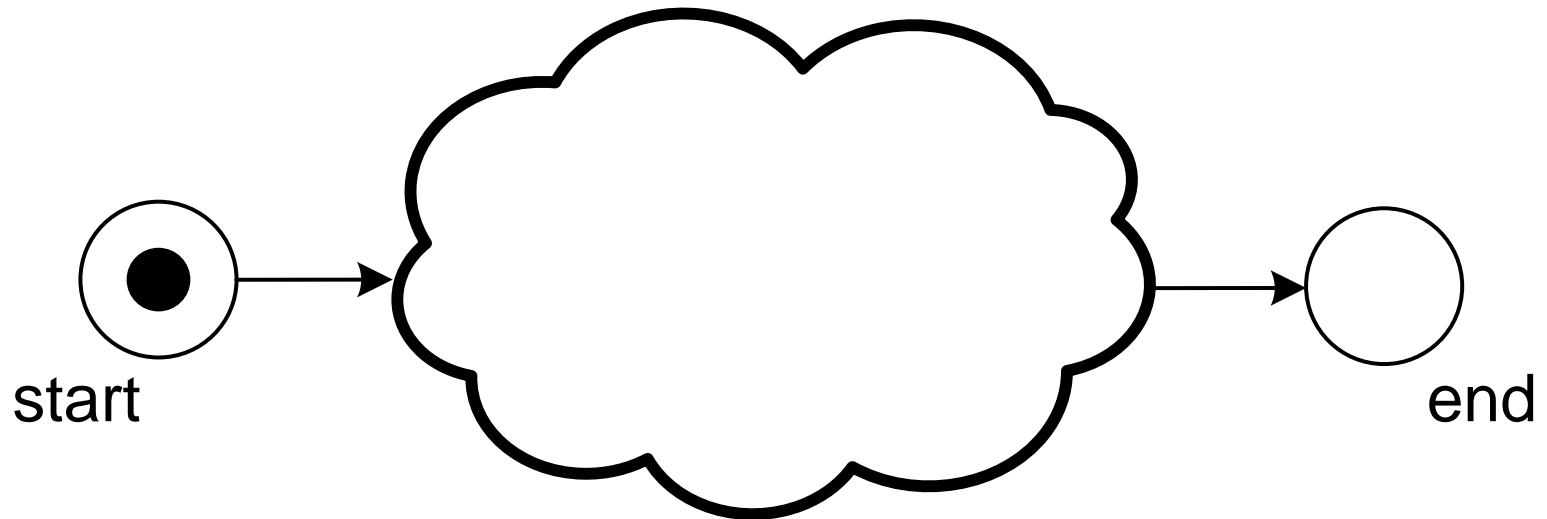


(b)

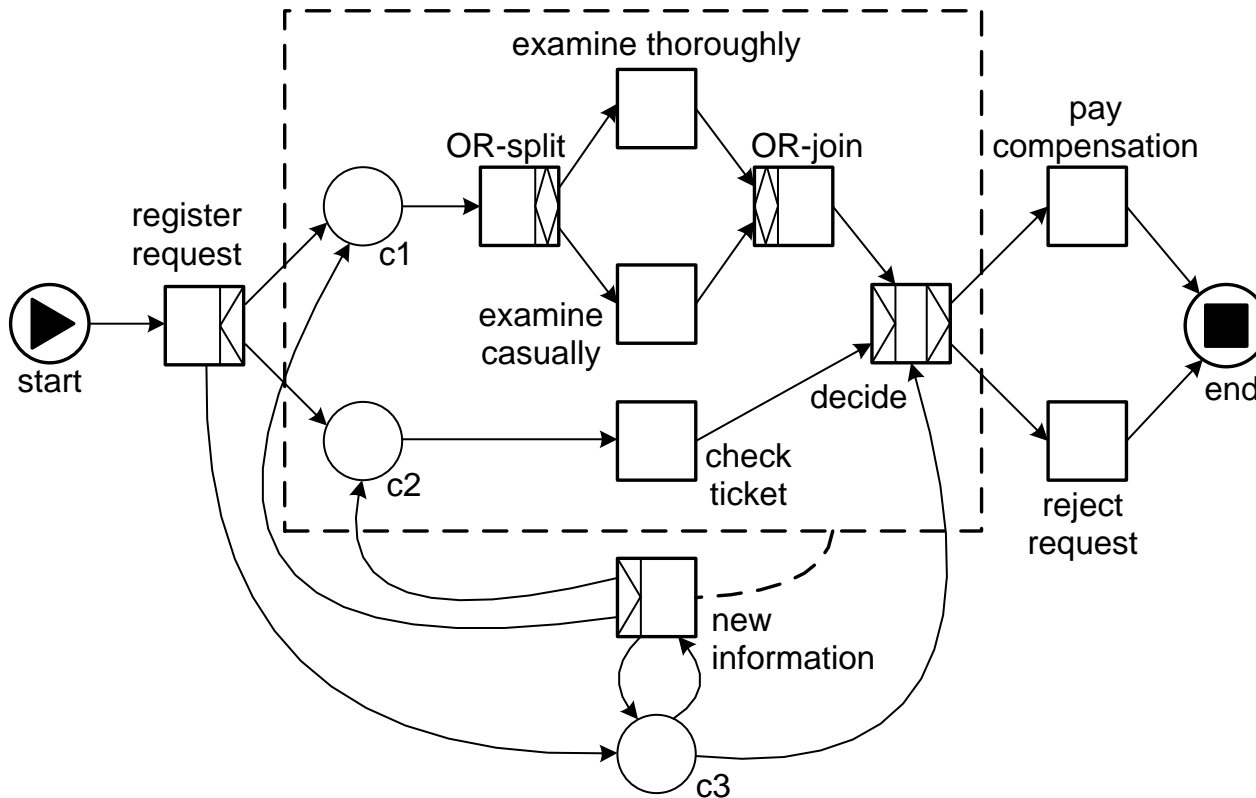
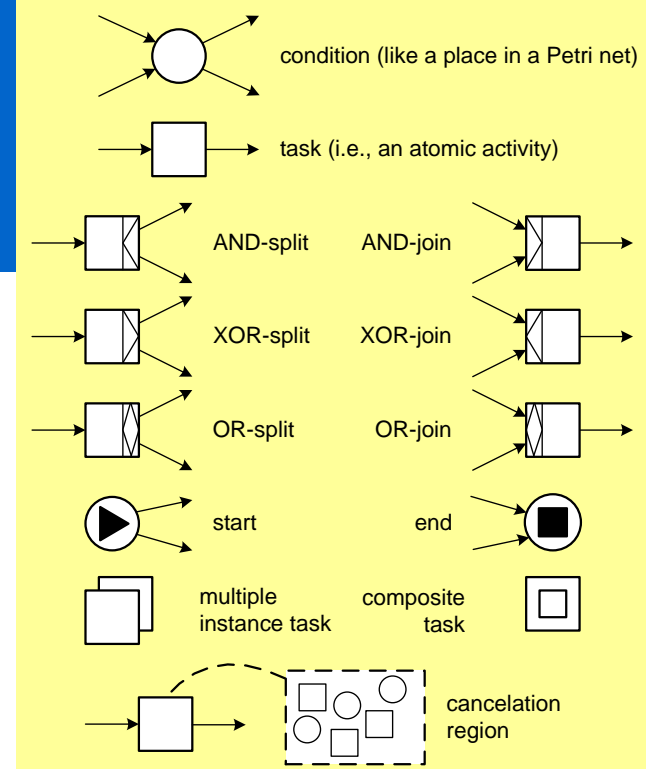


(c)

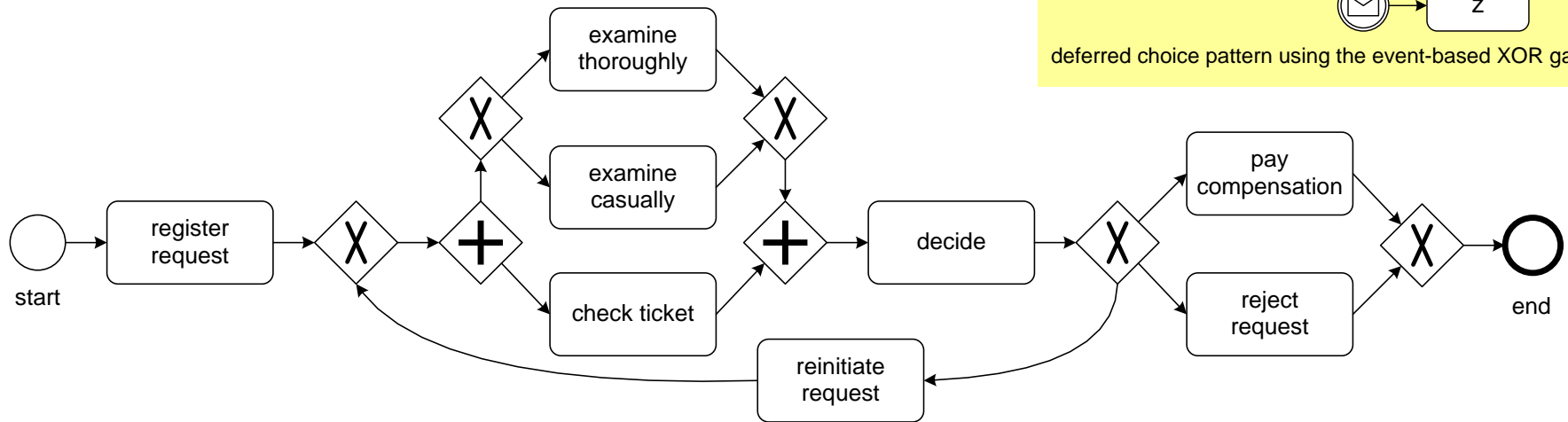
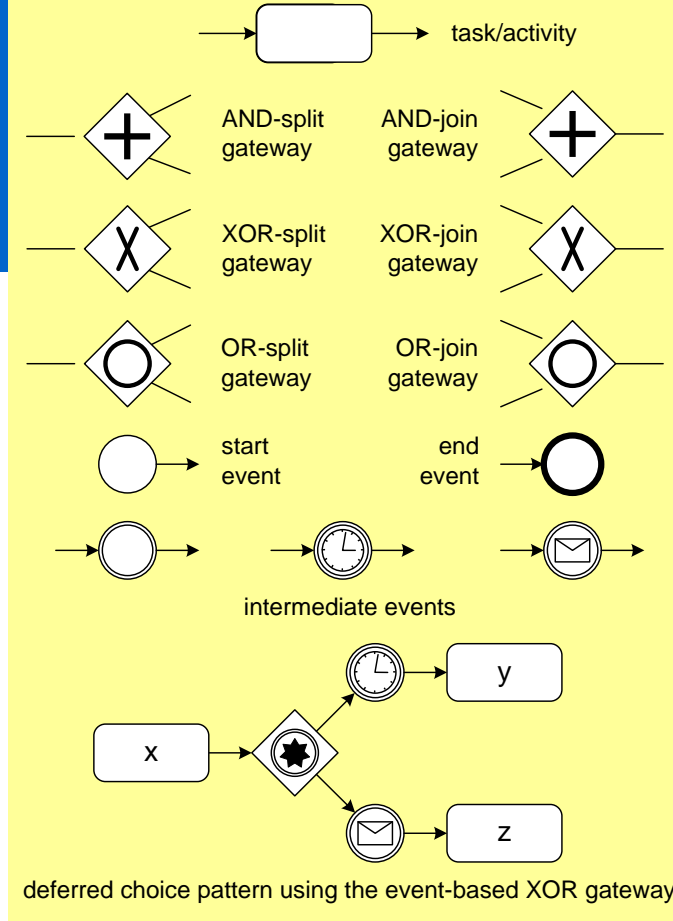
# WF-nets and soundness



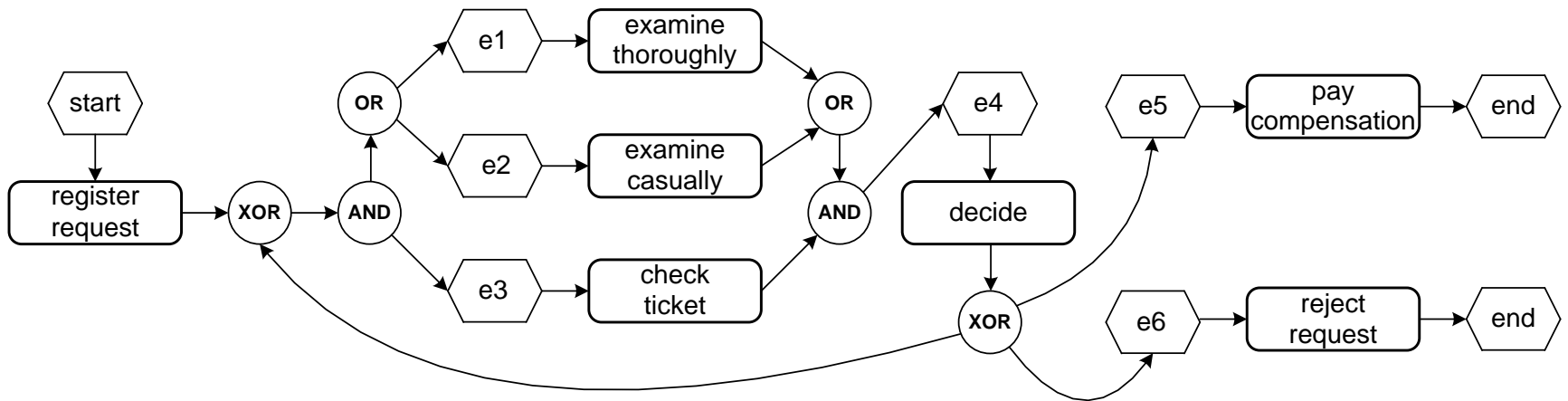
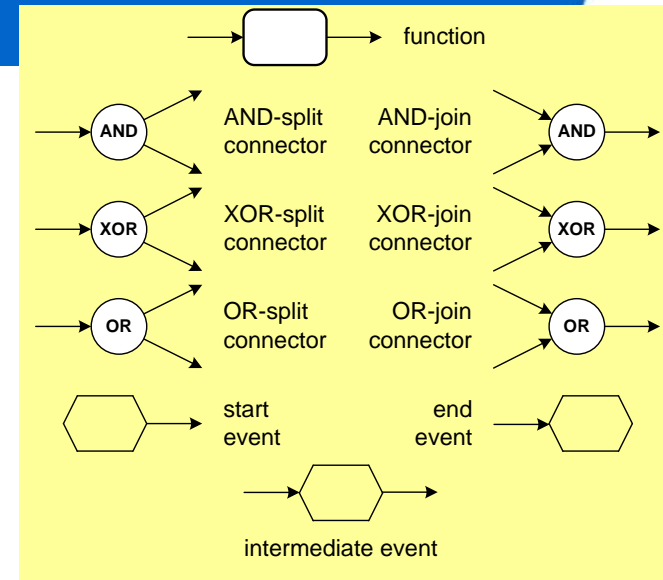
# YAWL



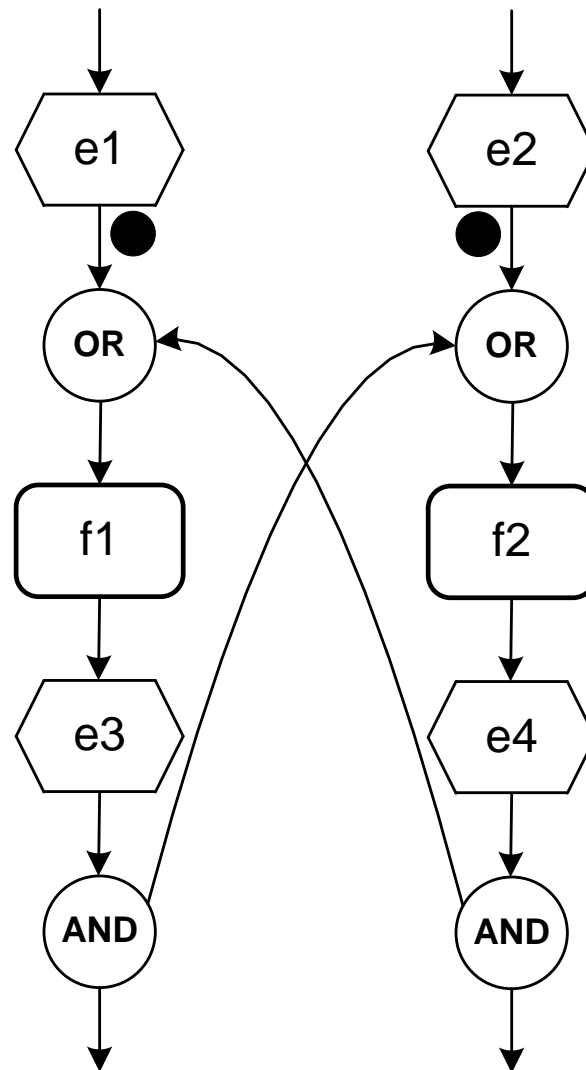
# BPMN



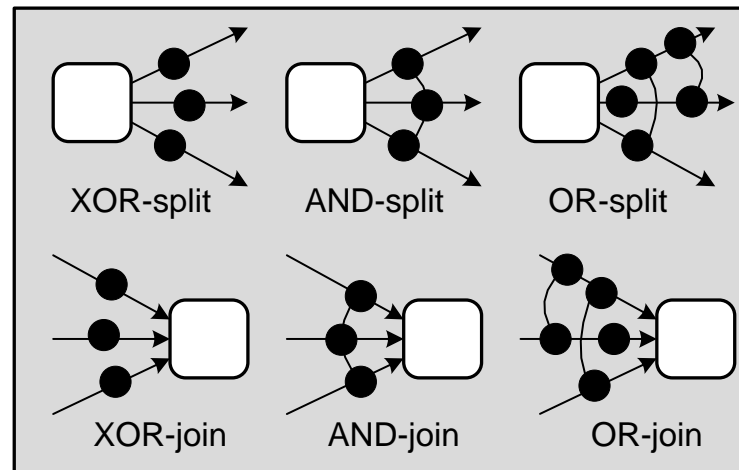
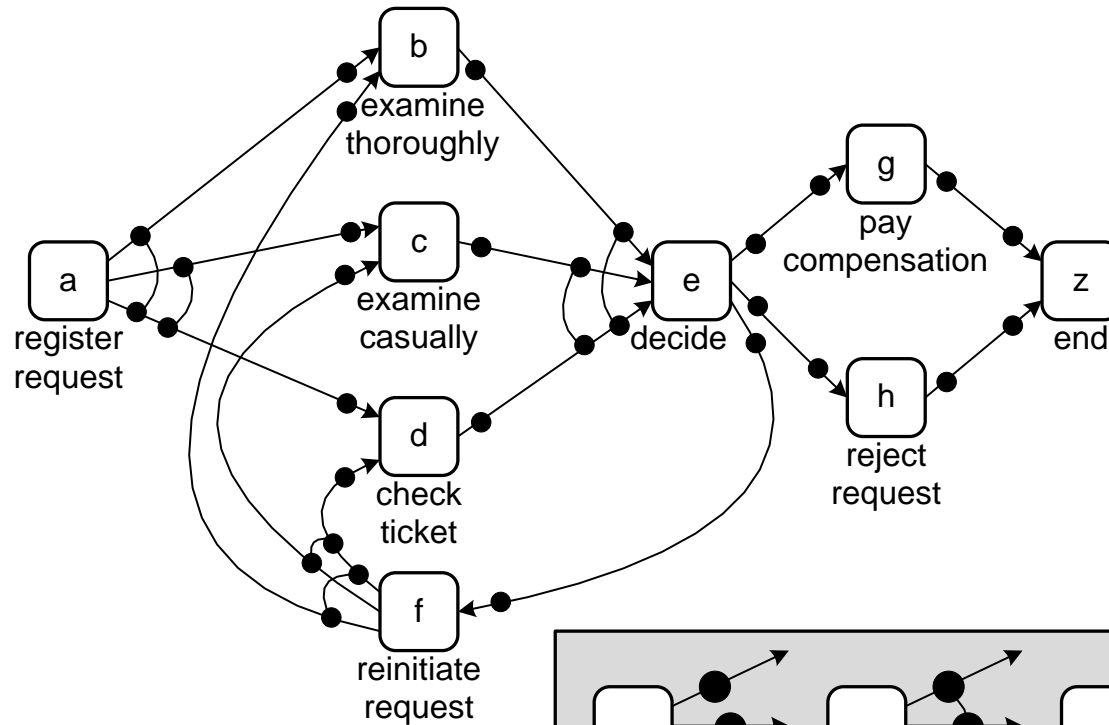
# Event-Driven Process Chains (EPCs)



# Vicious cycle paradox



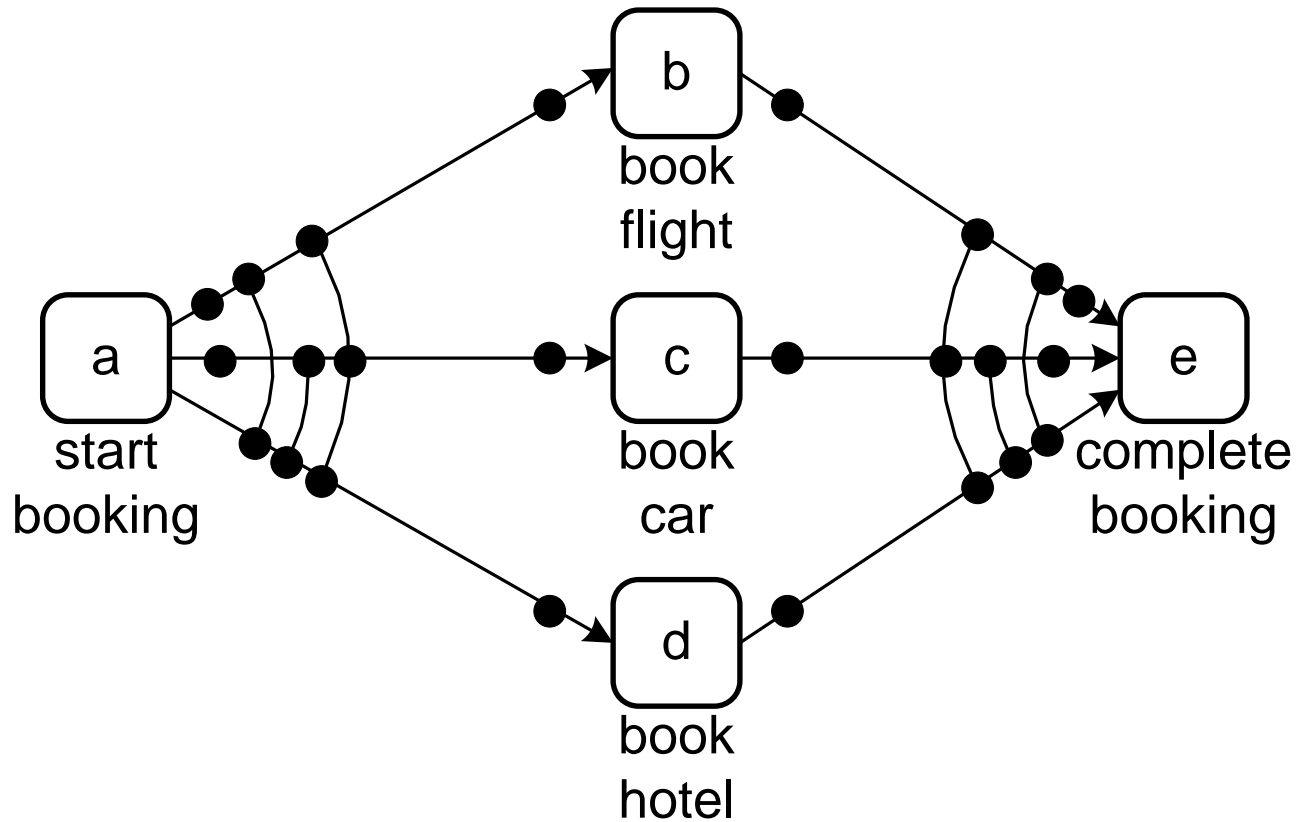
# Causal nets (C-nets)



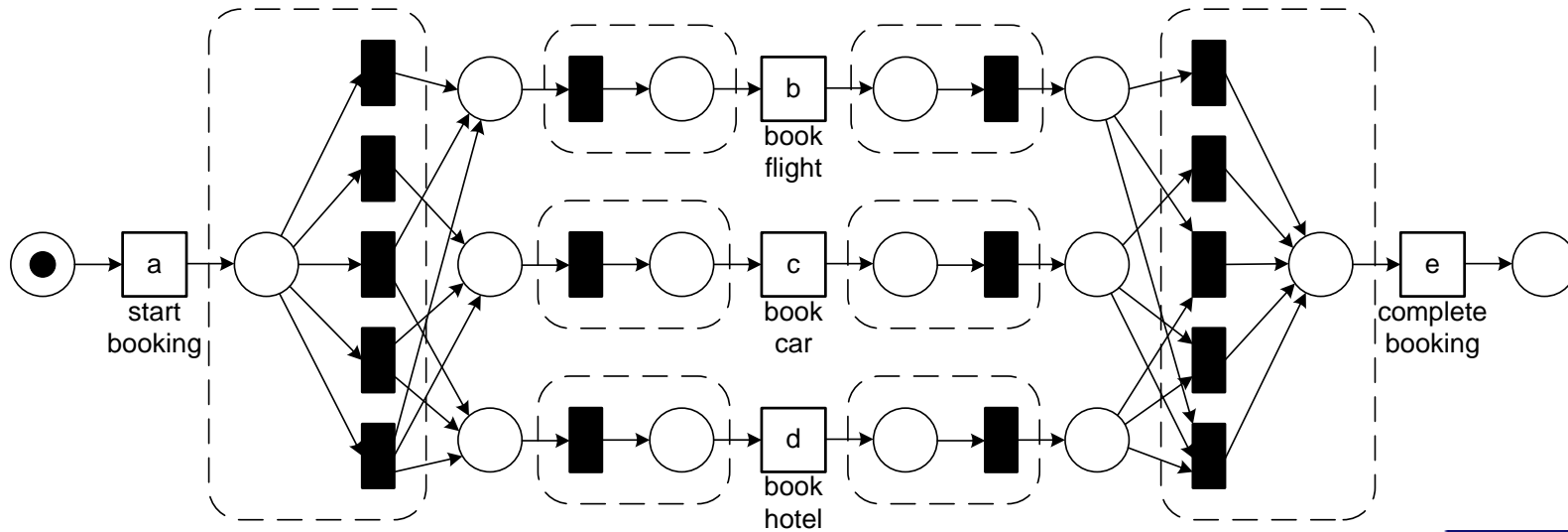
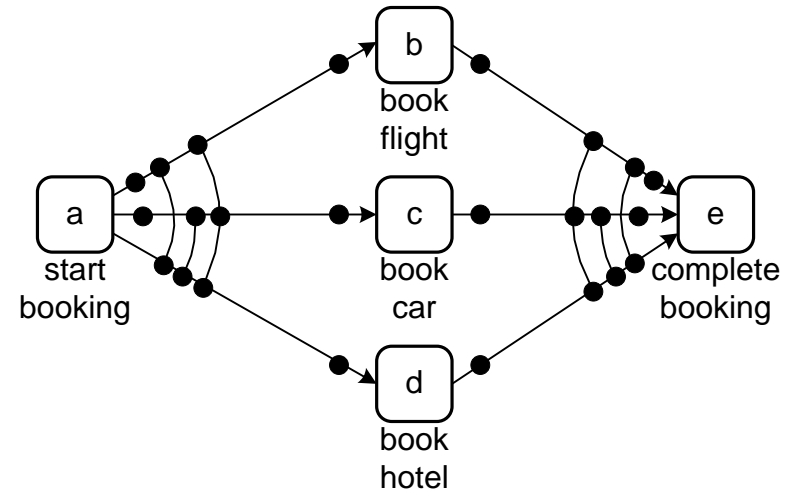
# Why C-nets?

- **Similar to heuristic nets and representation used by genetic miners.**
- **Fits well with mainstream languages (BPMN, EPCs, YAWL, BPEL, etc.).**
- **Model XOR, AND, and OR, but no silent steps or duplicate activities needed.**
- **Loose interpretation (focus on replay semantics rather than execution semantics).**

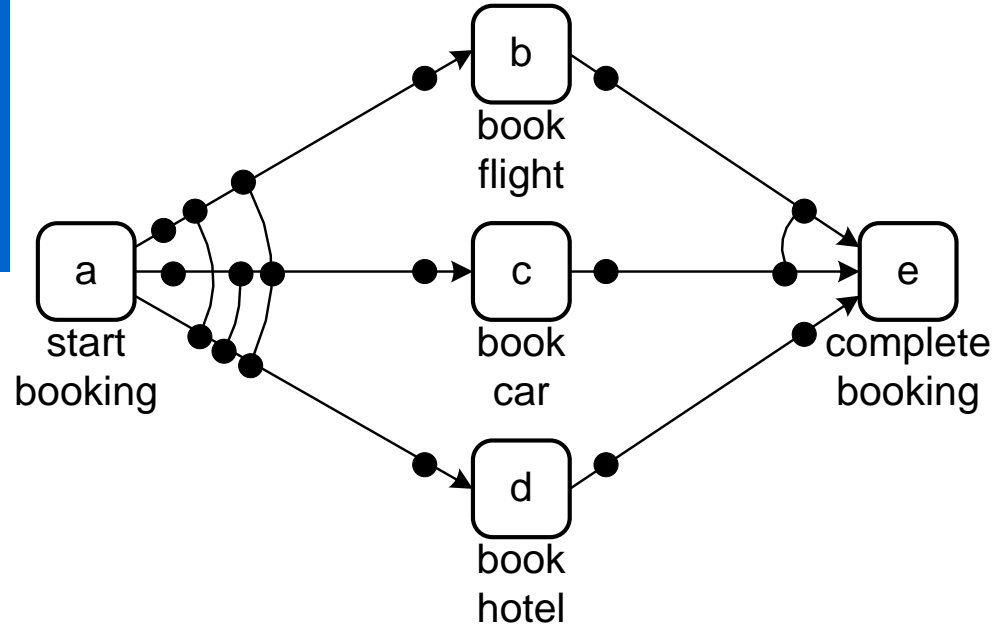
# Another C-net



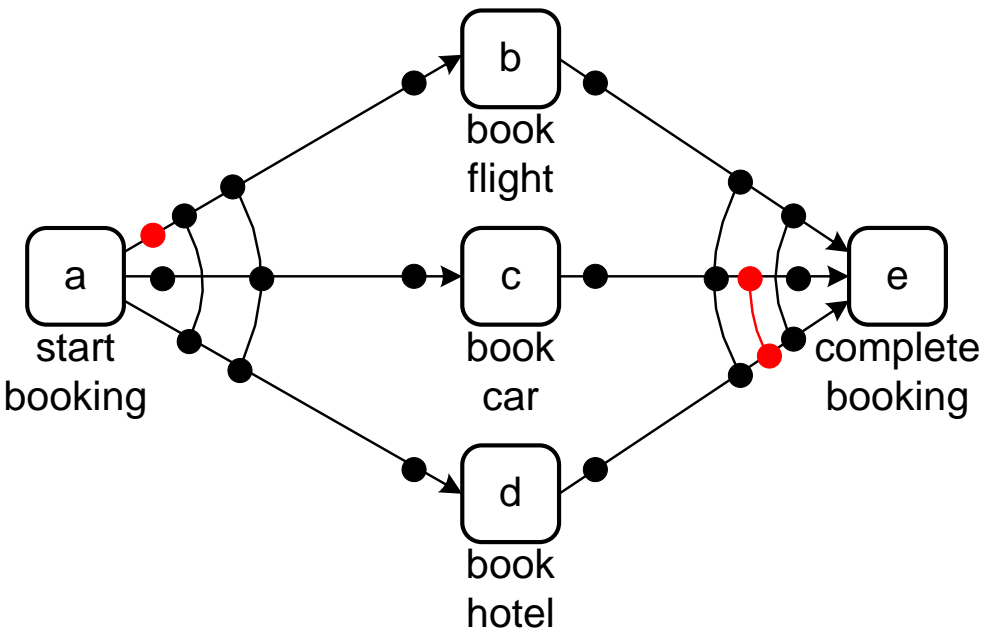
# WF-net interpretation of C-nets (only valid sequences!)



# Non-sound C-nets

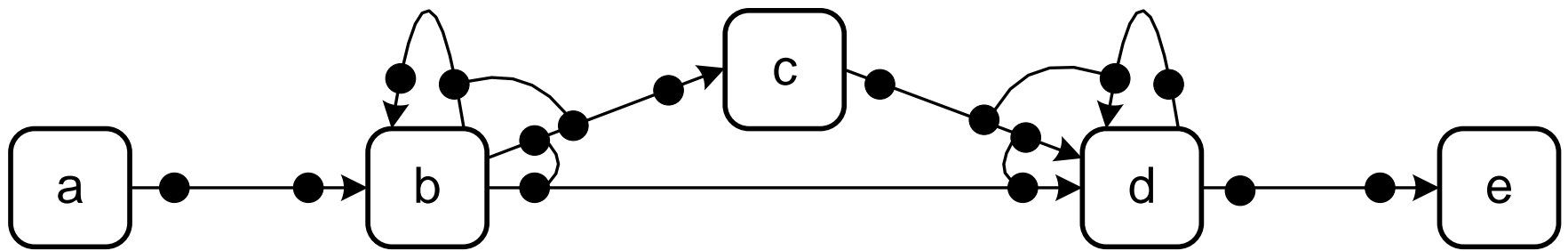


(a) unsound because there are no valid sequences

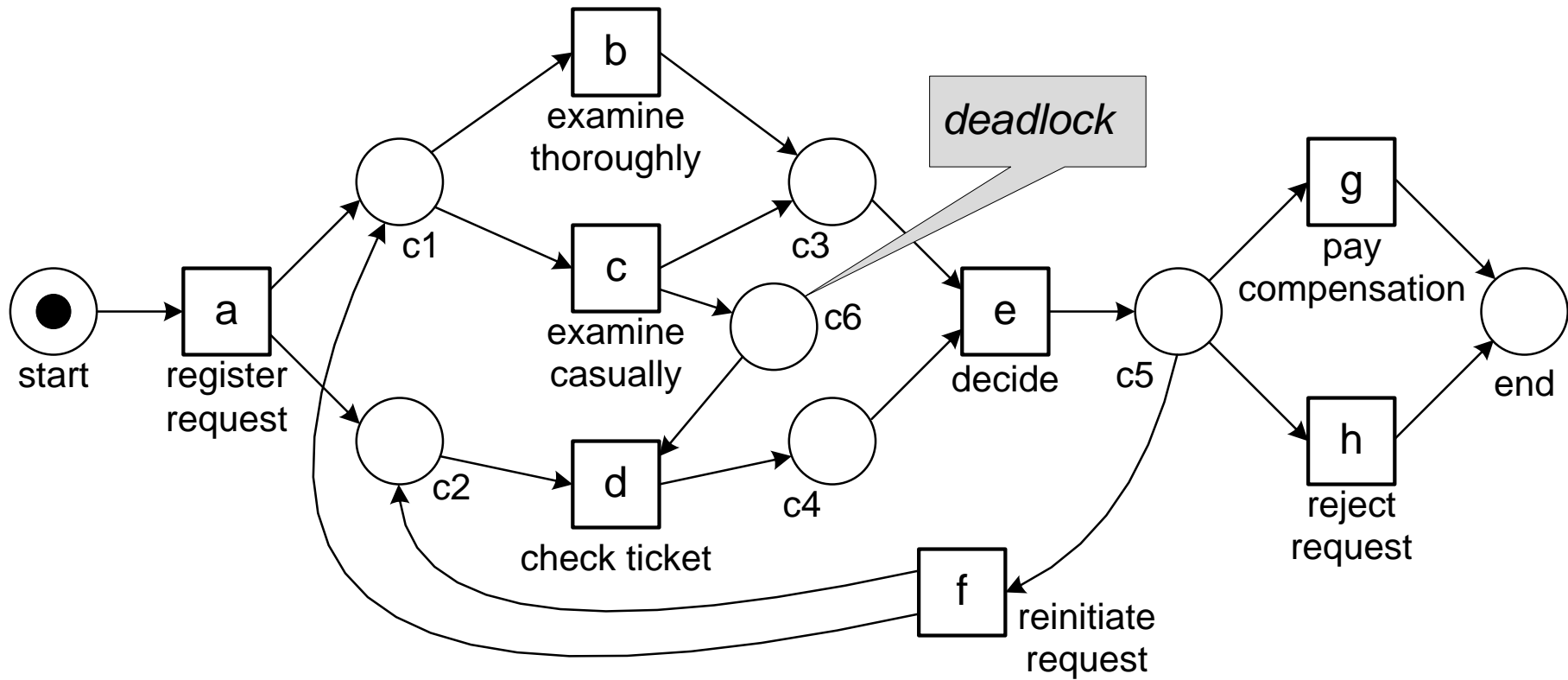


(b) unsound although there exist valid sequences

# Unbounded C-net



# Verification



# Example: YAWL

YAWLEditor - D:\application\_data\yawl\course 2010\insurance.yawl

Specification Net Edit Elements Tools View Help

Main\_Insurance\_Process

```
graph LR; start((start)) --> register[register]; register --> check_insurance[check insurance]; register --> check_damage[check damage]; check_insurance --> pay[pay]; check_damage --> c((c)); c --> dummy[dummy]; pay --> end((end)); dummy --> cancel[cancel]; cancel --> end;
```

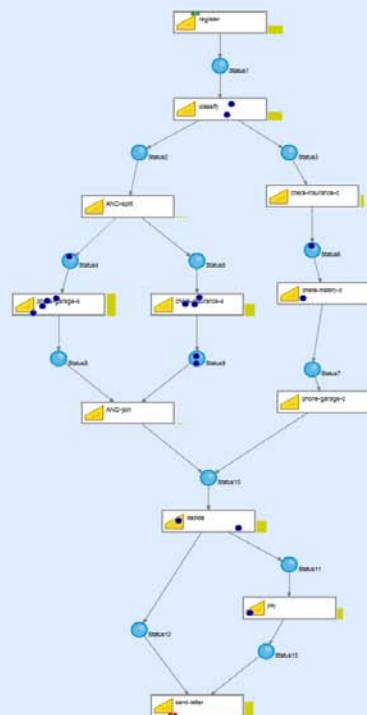
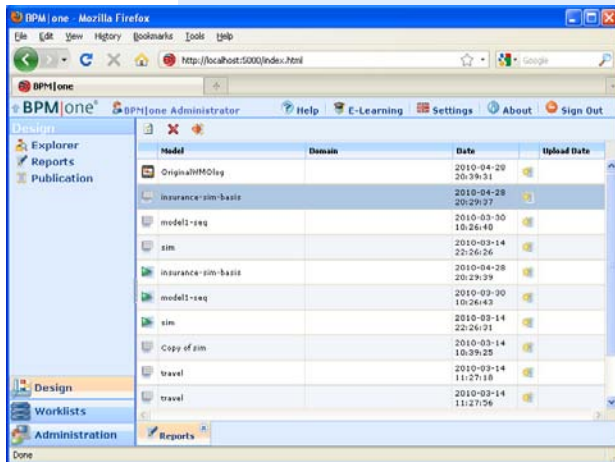
Task Icon  
No Icon  
Manual  
Automated  
Routing  
Plugin

ResetNet Analysis Warning: The net Main\_Insurance\_Process does not have an option to complete.  
ResetNet Analysis Warning: The net Main\_Insurance\_Process does not have proper completion. Markings larger than Mo can be found: 2OutputCondition\_2  
ResetNet Analysis Warning: The net Main\_Insurance\_Process does not satisfy the soundness property.

ResetNet Analysis Warning: The net Main\_Insurance\_Process does not satisfy the soundness property.

Use the palette toolbar to edit the selected net.

# Performance analysis, e.g., simulation in BPM|one



Roles										
	Mean	Lower 90%	Upper 90%	Lower 99%	Upper 99%	Mean	Lower 90%	Upper 90%	Lower 99%	Upper 99%
<<No role>>	0	0	0	0	0	0	0	0	0	0
employee	0.818765572	0.712002743	0.9255284	0.629159456	1.008371688					
Claim handler	0.477614884	0.381832289	0.573397479	0.307509182	0.647720586					
Claim handler A	0.87805118	0.598904828	0.761197533	0.53316285	0.824898911					

Activities										
	Mean	Lower 90%	Upper 90%	Lower 99%	Upper 99%	Mean	Lower 90%	Upper 90%	Lower 99%	Upper 99%
register	5.683619366	0.750153688	10.61708544	-3.078000163	14.44523929	10.12389217	9.68612823	10.55068811	9.364586531	10.8821978
phone-garage-s	2.888949042	0.587130287	5.190767798	-1.188980442	6.976878526	29.95317042	29.69568701	30.21065383	29.49589115	30.4104496
check-insurance-s	1.741425234	0.724631759	2.758218708	-0.064355527	3.547205994	14.85471567	14.58402909	15.12540225	14.37398814	15.3354443
check-history-c	1.027479575	-0.591329722	2.646288871	-1.847454968	3.902414118	19.80467549	19.20910654	20.40024445	18.74697106	20.8623799
AND-join	0	0	0	0	0	0	0	0	0	0
phone-garage-c	0.758414809	-0.021672562	1.53850218	-0.626986261	2.145815878	29.97848491	29.6740688	30.34995294	29.31946659	30.6375032
decide	1.687003061	1.046399418	2.327606703	0.549318996	2.824687125	15.24293539	15.04371202	15.44215877	14.8891234	15.5967473
classify	1.570212696	0.667263861	2.473188331	-0.033395359	3.173847551	9.87474254	9.696300158	10.05318492	9.55783667	10.1916484
pay	2.146984462	0.656782494	3.63718643	-0.499549088	4.793518011	15.17529939	14.73446303	15.61613575	14.39239329	15.9582054
send-letter	1.150684282	-0.028173015	2.327485578	-0.939841103	3.240653867	20.29803097	20.022110823	20.49499337	19.83726853	20.8787953
AND-split	0	0	0	0	0	0	0	0	0	0
check-insurance-c	0.877120927	0.221961283	1.532280572	-0.286413961	2.040655815	14.84699658	14.12850766	15.56542549	13.87101495	16.122918

Status										
	Mean	Lower 90%	Upper 90%	Lower 99%	Upper 99%	Mean	Lower 90%	Upper 90%	Lower 99%	Upper 99%
Status1	0	0	0	0	0	1.570226096	0.667263861	2.473188331	-0.033395359	3.17384755
Status8	0.013955404	-0.011582986	0.039493794	-0.031399966	0.059310468	0.013955404	-0.011582986	0.039493794	-0.031399966	0.059310468
Status9	15.76128093	13.74211032	17.78045154	12.17532219	19.84723967	15.76128093	13.74211032	17.78045154	12.17532219	19.84723967
Status7	0	0	0	0	0	0.758414809	-0.021672562	1.53850218	-0.626986261	2.145815878
Status10	0	0	0	0	0	1.687003061	1.046399418	2.327606703	0.549318996	2.824687125
Status11	0	0	0	0	0	2.146984462	0.656782494	3.63718643	-0.499549088	4.793518011
Status12	0	0	0	0	0	1.177037517	-0.240102843	2.994177877	-1.494930368	4.24900810
Status15	0	0	0	0	0	0.923883831	0.081404857	1.766271806	-0.572286892	2.41996335
Status2	0	0	0	0	0	0	0	0	0	0
Status3	0	0	0	0	0	0.877120927	0.221961283	1.532280572	-0.286413961	2.040655815
Status4	0	0	0	0	0	2.888949042	0.587130287	5.190767798	-1.188980442	6.976878526
Status5	0	0	0	0	0	1.741425234	0.724631759	2.758218708	-0.064355527	3.547205994
Status6	0	0	0	0	0	1.027479575	-0.591329722	2.646288871	-1.847454968	3.902414118

Total										
	Mean	Lower 90%	Upper 90%	Lower 99%	Upper 99%	Mean	Lower 90%	Upper 90%	Lower 99%	Upper 99%
Lead time	115.6947846	105.258757	126.1308123	97.16085576	134.2287135	115.4428918	110.4228233	116.4629603	108.0793822	118.8064014
Work time										

11-05-2010 13:00

# Limitations of model-based analysis

- **Verification and performance analysis heavily rely on the availability of high quality models.**
- **When the models and reality have little in common, model-based analysis does not make much sense.**
- **There is often a lack of alignment between hand-made models and reality**
- **Process mining aims to address these problems by establishing a direct connection between the models and actual low-level event data about the process.**
- **Process discovery techniques allow for viewing the same reality from different angles and at different levels of abstraction.**