

Process Mining and Its application to telecom systems

Behzad Bordbar

School of Computer Science,
University of Birmingham, UK



Process Mining



New title



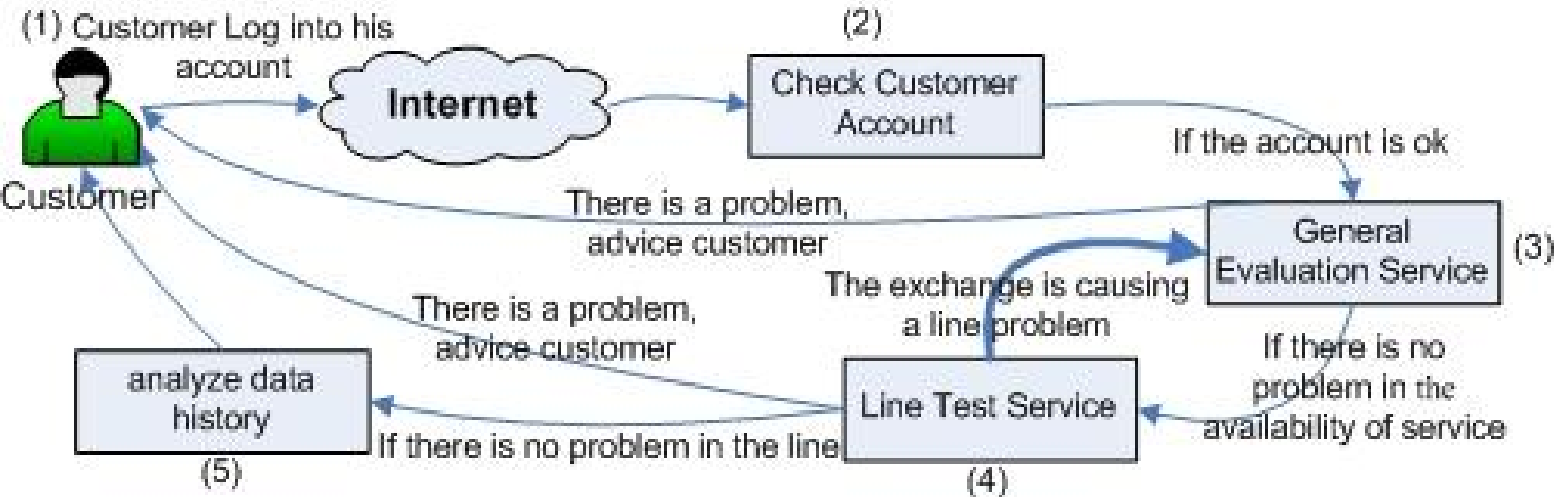
Outline

- ❑ Service oriented Architectures
- ❑ Process Mining
- ❑ Why process mining?
- ❑ Current research in Birmingham
- ❑ Application of PM to Multi-disciplinary research (Shail)



A typical SoA

Service oriented Architecture (SoA)

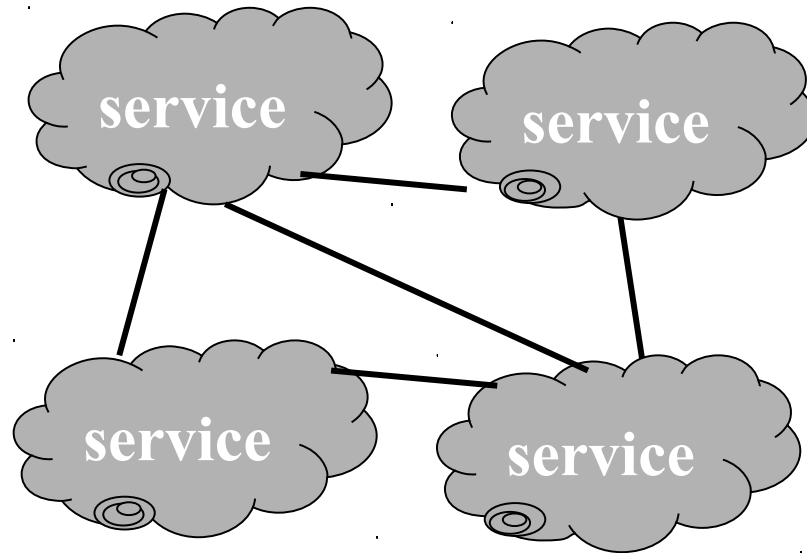


Example2: booking a holiday (airline, hotel, car hire, google map...)

Services interacting with each other



An abstract view



Question: Given the same fixed structure
Which service starts first?

We need information about the **dynamics!**

What behaviour are sequential, parallel, repetitive...

Business Process



Tool for BP design and execution:

Software tools

1) Design BP

2) Deploy them

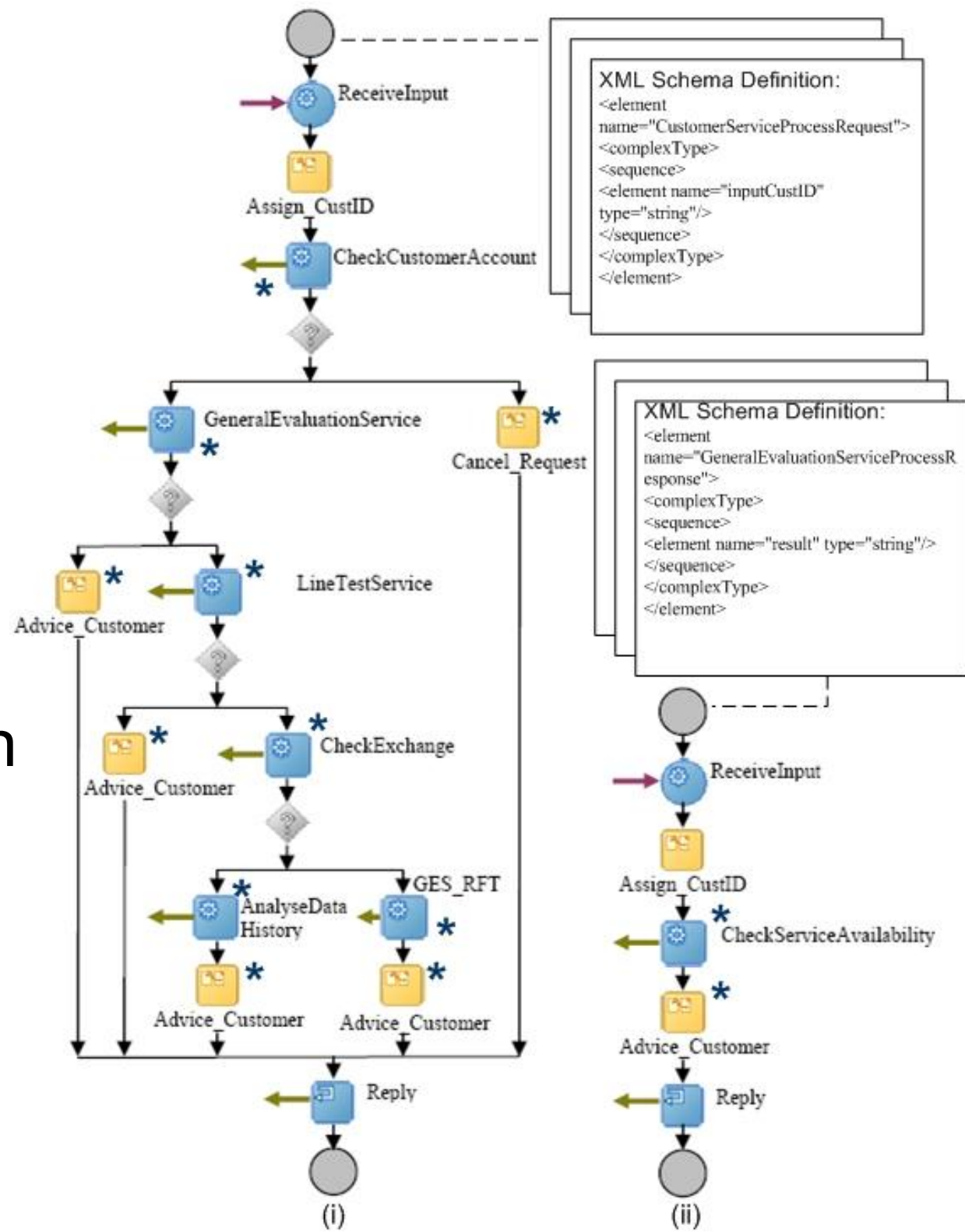
Tools takes care of the code

Generation, execution

.... job for different

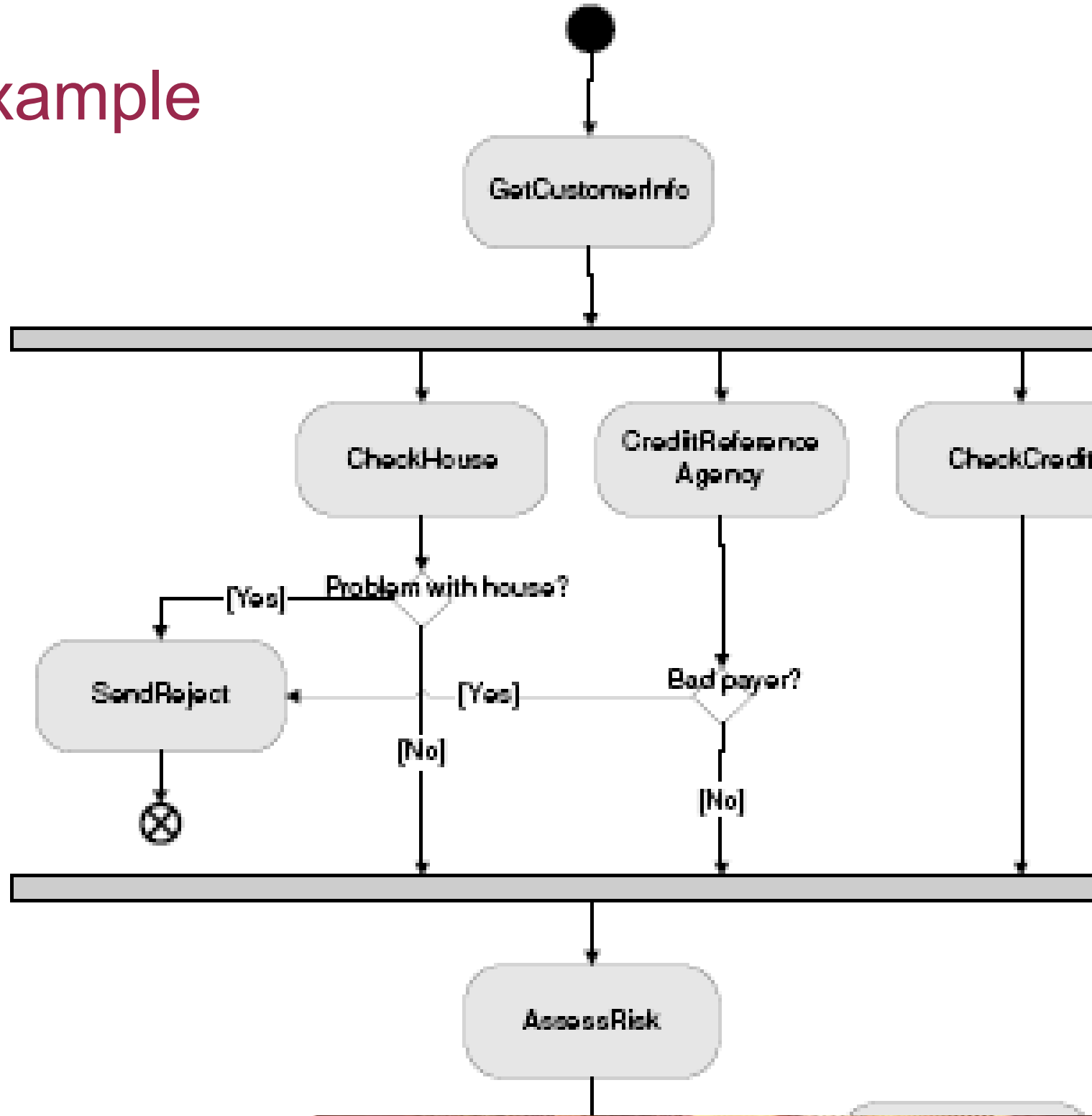
users separated!

Very complex Engineering!



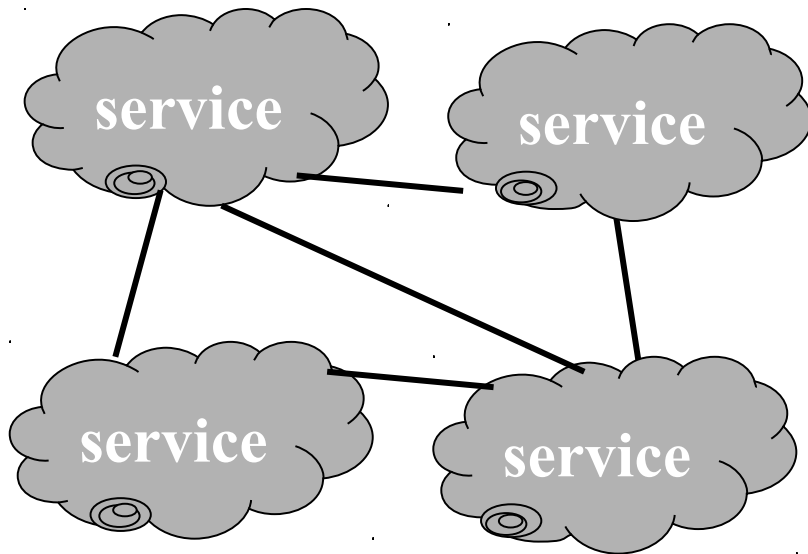
A simplified example Mortgage evaluation

Notice:
a correct
BP is
well-structured.



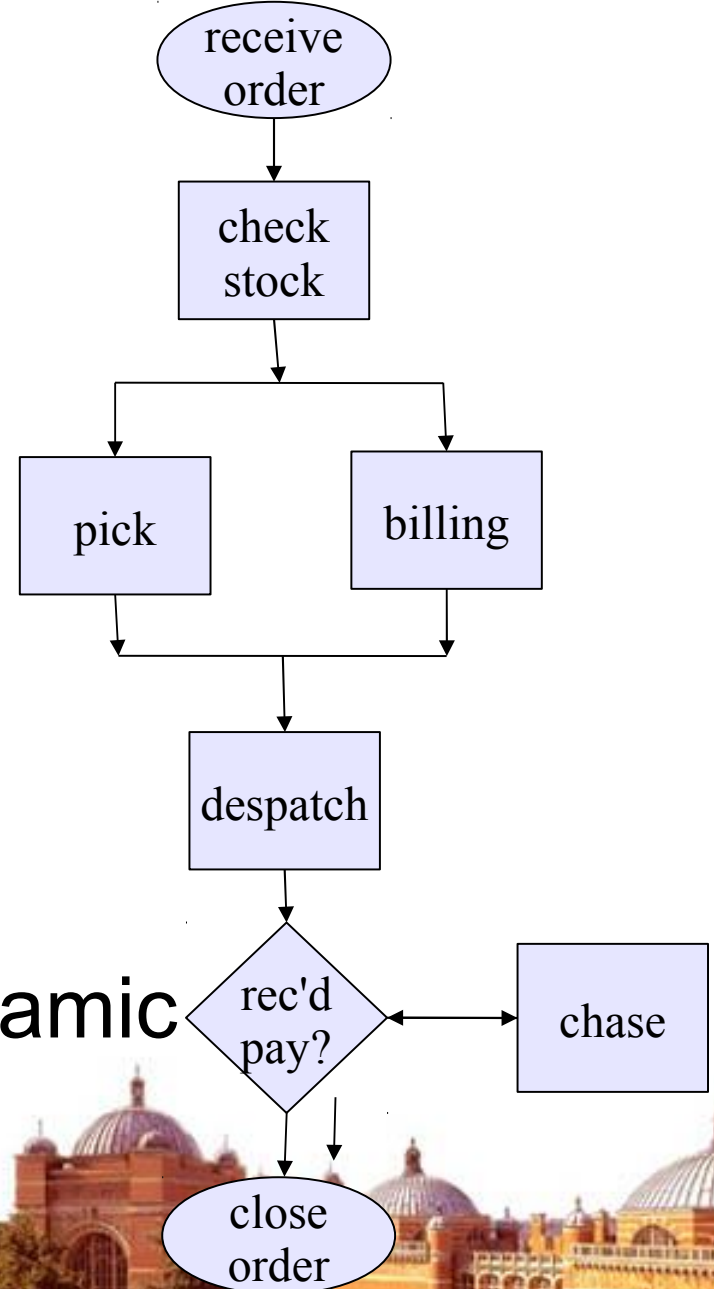
An abstract view (2)

Business Process

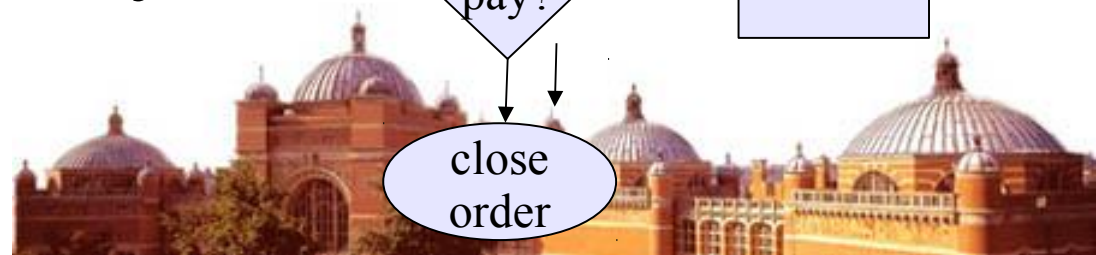


Static

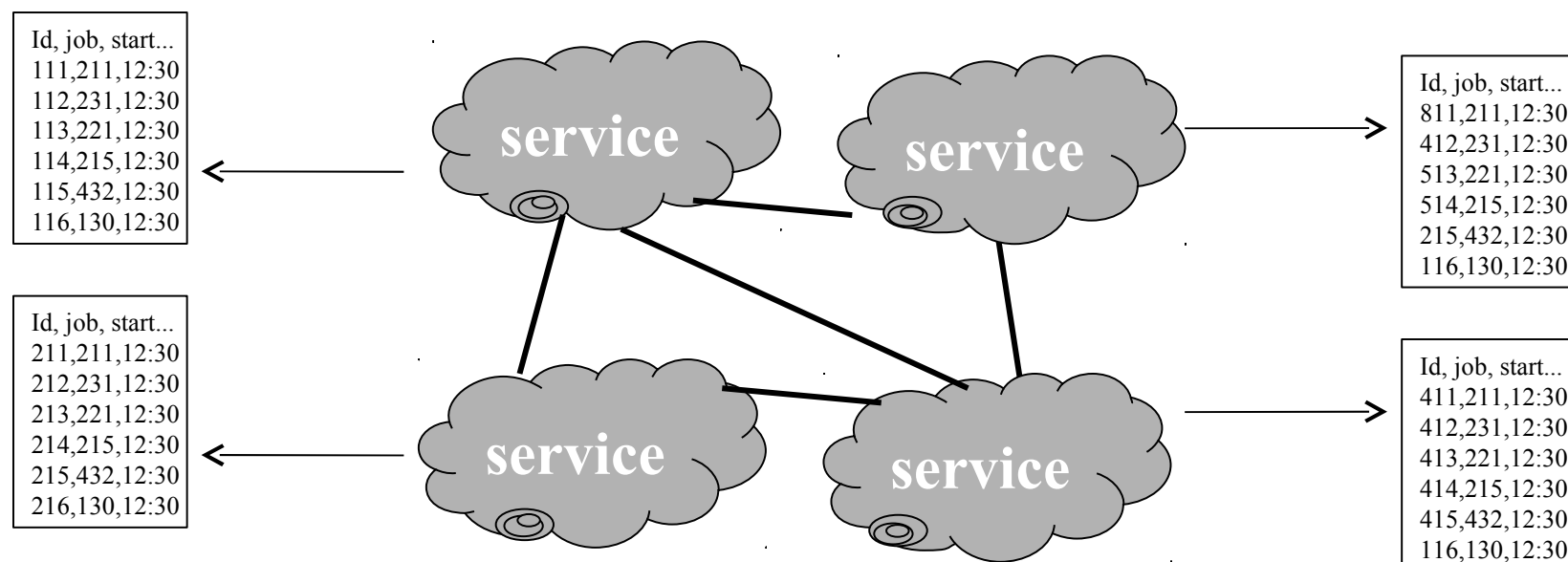
plus



Dynamic



Execution of the BP creates logs on each server



Log: id+ job+start time +end time+....

Use id to identify sequence of task for each user

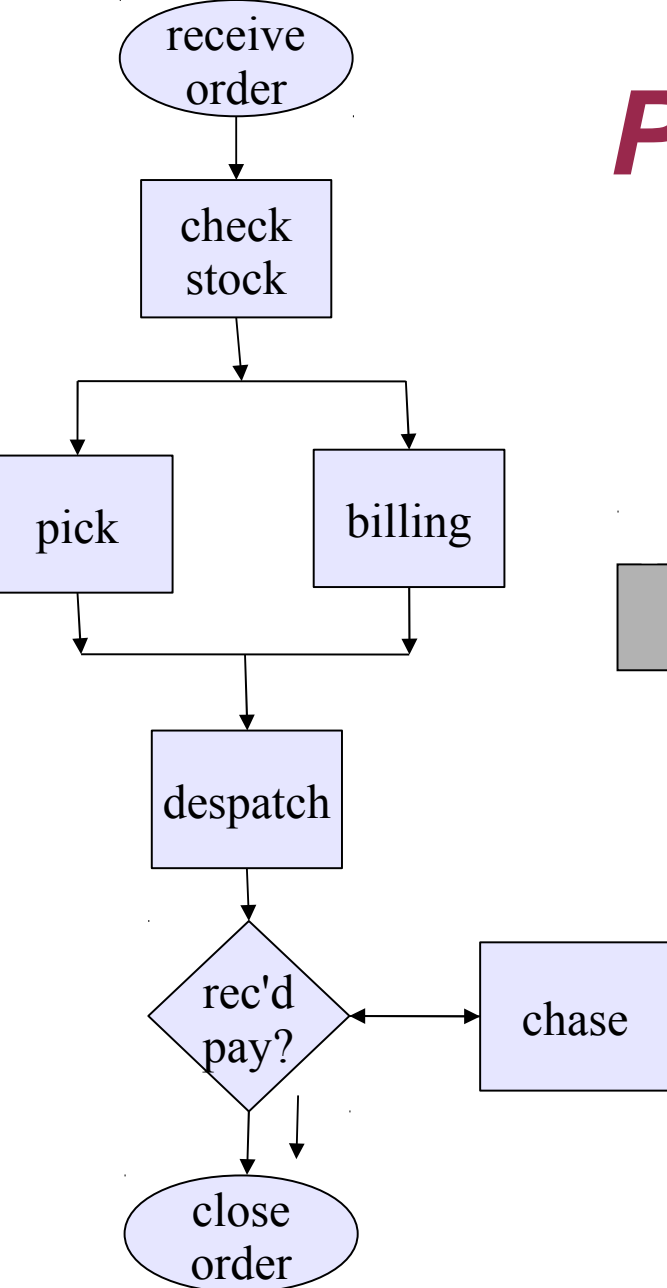
Using time Logs can (often) be integrated

Process mining extracts a “True” BP model from log files

True Business Process? (people, fault, bad design, ..)

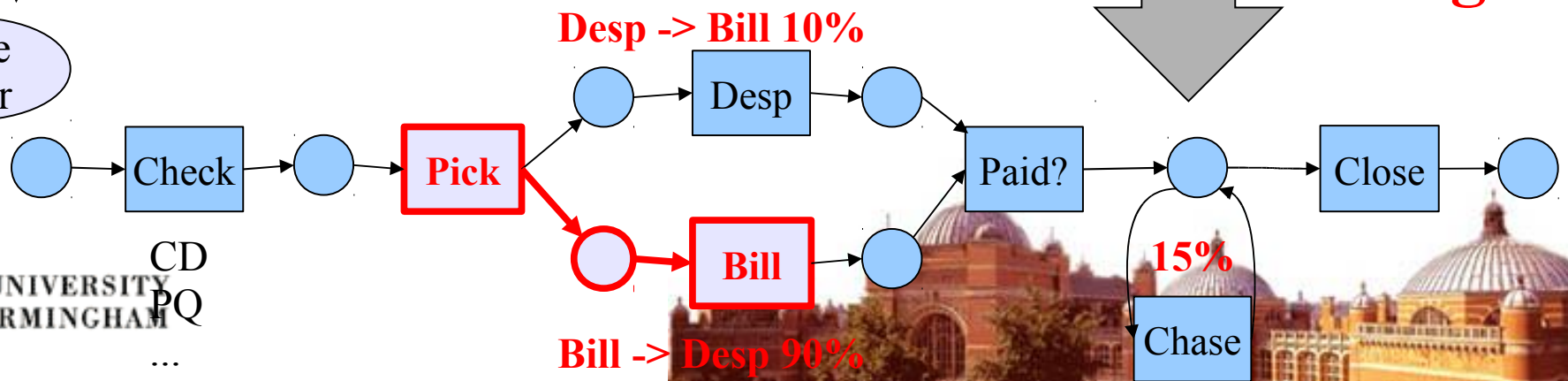


Process Mining



Date	Case ID	User	Task	Other Data...
20100714	0001	AB	Rec	orderno
20100714	0001	CD	Check	---
...	0001	XY	Pick	---
...	0002	AB	Rec	orderno
...	0001	MN	Billing	BACSxxxx
...	0002	PQ		fail
...	0003	AB		orderno

Process Mining



Central to PM: algorithms

Algorithms (20+)

- Alpha, Alpha++ (formal)
- HeuristicsMiner, Genetic Miner, Region Miner (practical)
- Probabilistic Approaches (Datta, Herbst)
- Clustering (Variants)
- Fuzzy (Abstraction of complex or flexible processes)
- ...



Example: alpha alg.

Given two tasks a and b, is there any log file with a followed by b **immediately**?

written as $(a>b)$

- ◆ $a>b$ and $b>a$ then $a||b$
- ◆ $a>b$ and never $b>a$ then $a \rightarrow b$
- ◆ Never $a>b$ and never $b>a$ then $a \#b$

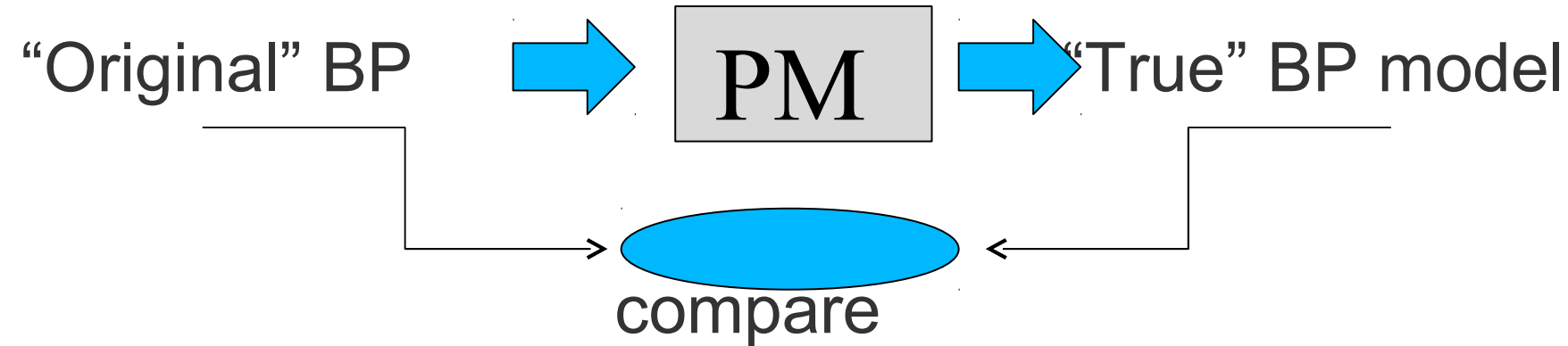
... then put everything together and create the “True” BP

Problem: what if 100000 times $a >b$ and one $b>a$?

Noise or a rare case?



Why Process Mining?



- ❑ Troubleshooting
 - ❑ **Why** is the model different?
 - ❑ **Fault/error** in implementation
- ❑ Streamlining
 - ❑ Order and **timing** of event
 - ❑ where are the bottlenecks? **Load balancing**
 - ❑ **People** interactions: is work passed efficiently? (provide training!)
- ❑ Audit and governance (**emerging area**)
 - ❑ Show **conformance** to SLA, lack of security violations,...
- ❑ Planning



Our current research

For a given problem, some algorithm perform better than others

Problem 1: How to compare PM algorithms?

Aim: Presenting guideline for choosing suitable algorithm for a given problem

Problem 2: is it possible to use PM in real-time or near-real-time? (to avoid undesirable situations)

Aim: If a percentage of the log files (instead of all) is used, what is the probability of identifying the “True” BP model accurately? (approximation problem)

Collaboration with **Peter Tino** and BT research team.



Application of PM to Multi-disciplinary research (Shail)

1. Data mining has been used successfully at snapshots of data, is it possible to use PM techniques to apply data mining in a temporal manner?
2. Distributed PM seems to provide a solution to anonymity (major issue).

Formulation: is it possible to design a protocol (alg) so that PM on the nodes can produce the same results as PM on the integrated data?

Observation: evolutionary techniques are underused in PM, while being successfully used in Data mining.



Summary

- ❑ Brief introduction to BP and PM
- ❑ Some of the applications of PM
- ❑ Current research at B'ham
- ❑ Application of PM

