

# What can happen? Semantic Analysis of Corporate Real-Time Data Stream

**Kia Teymourian**

**Corporate Semantic Web Research Group**

Freie Universität Berlin

Innovationsforum Semantic Media Web 2014

Berlin,

7 October, 2014



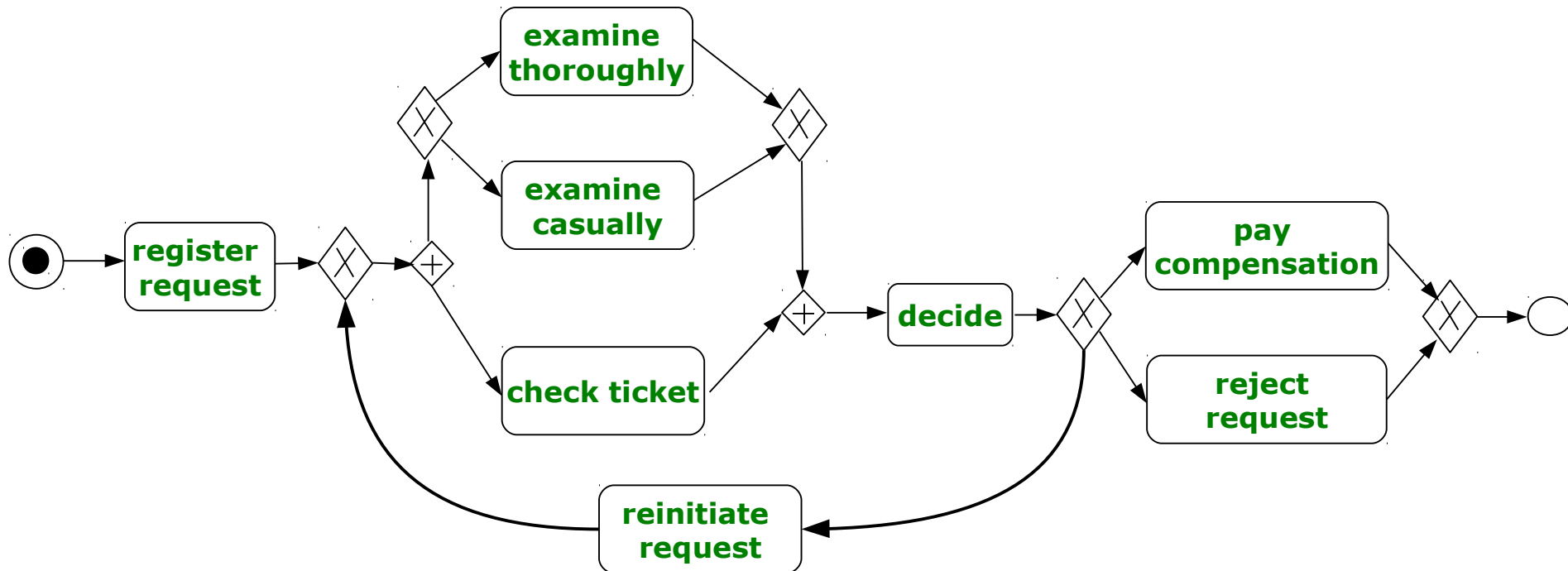
# Outline

- Activity Logs in Business Processes
- Mining of Complex Patterns from Logs
- Knowledge-based Event Patterns
- Semantic Analysis of Event Logs

# Motivation

- In today's corporations we have huge amounts of:
  - **Tracking Data (logs, historical data)**
  - **Streaming Data (real-time data)**
  - **Metadata**
- **Monitoring** of Business Processes
  - Business Process Mining (Discovery, Conformance and Extension)
- Detection of **Complex Events**
  - specification of complex event patterns
- **Mining** of Complex Event Pattern

# Example Process



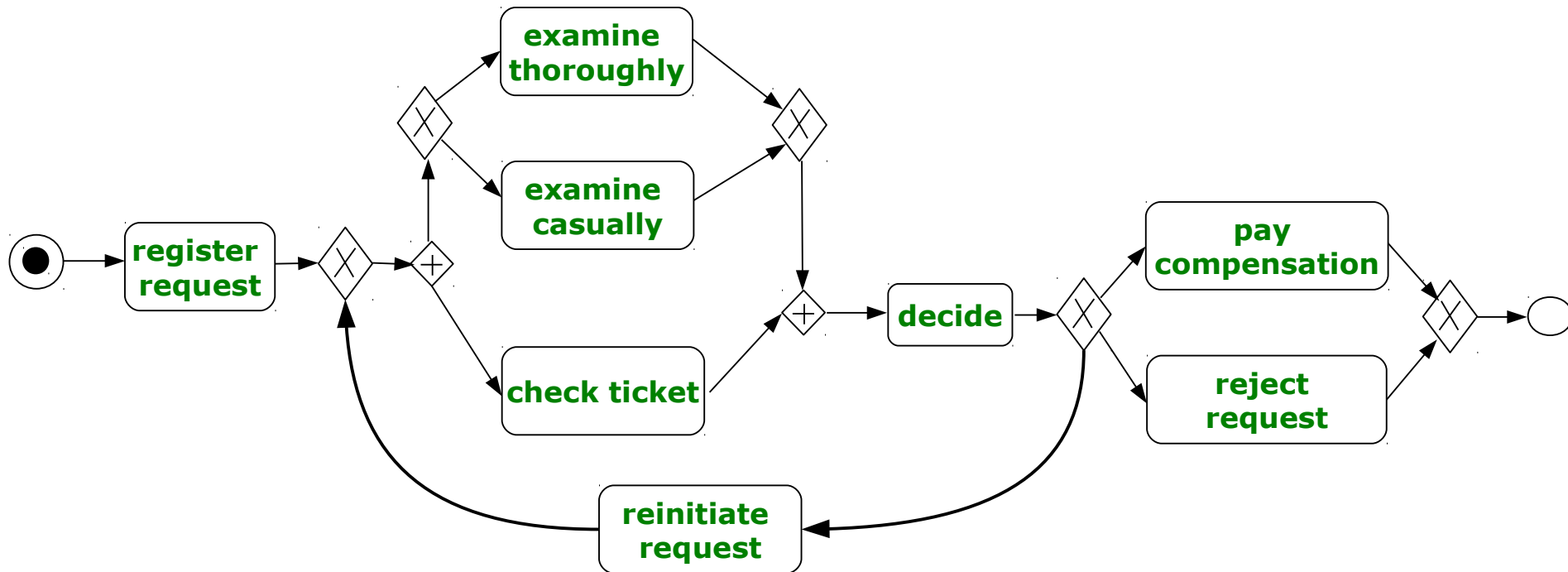
- BPMN Process for **Request for Compensation within an Airline**

# Example Event Log

Case id	Event id	Properties				
		Timestamp	Activity	Resource	Cost	...
1	35654423	30-12-2010:11.02	Register request	Pete	50	...
	35654424	31-12-2010:10.06	Examine thoroughly	Sue	400	...
	35654425	05-01-2011:15.12	Check ticket	Mike	100	...
	35654426	06-01-2011:11.18	Decide	Sara	200	...
	35654427	07-01-2011:14.24	Reject request	Pete	200	...
2	35654483	30-12-2010:11.32	Register request	Mike	50	...
	35654485	30-12-2010:12.12	Check ticket	Mike	100	...
	35654487	30-12-2010:14.16	Examine casually	Pete	400	...
	35654488	05-01-2011:11.22	Decide	Sara	200	...
	35654489	08-01-2011:12.05	Pay compensation	Ellen	200	...
3	35654521	30-12-2010:14.32	Register request	Pete	50	...
	35654522	30-12-2010:15.06	Examine casually	Mike	400	...
	35654524	30-12-2010:16.34	Check ticket	Ellen	100	...
	35654525	06-01-2011:09.18	Decide	Sara	200	...
	35654526	06-01-2011:12.18	Reinitiate request	Sara	200	...
	35654527	06-01-2011:13.06	Examine thoroughly	Sean	400	...
	35654530	08-01-2011:11.43	Check ticket	Pete	100	...
	35654531	09-01-2011:09.55	Decide	Sara	200	...
35654533	15-01-2011:10.45	Pay compensation	Ellen	200	...	
4	35654641	06-01-2011:15.02	Register request	Pete	50	...

Example adapted from  
(van der Aalst et al., 2011)

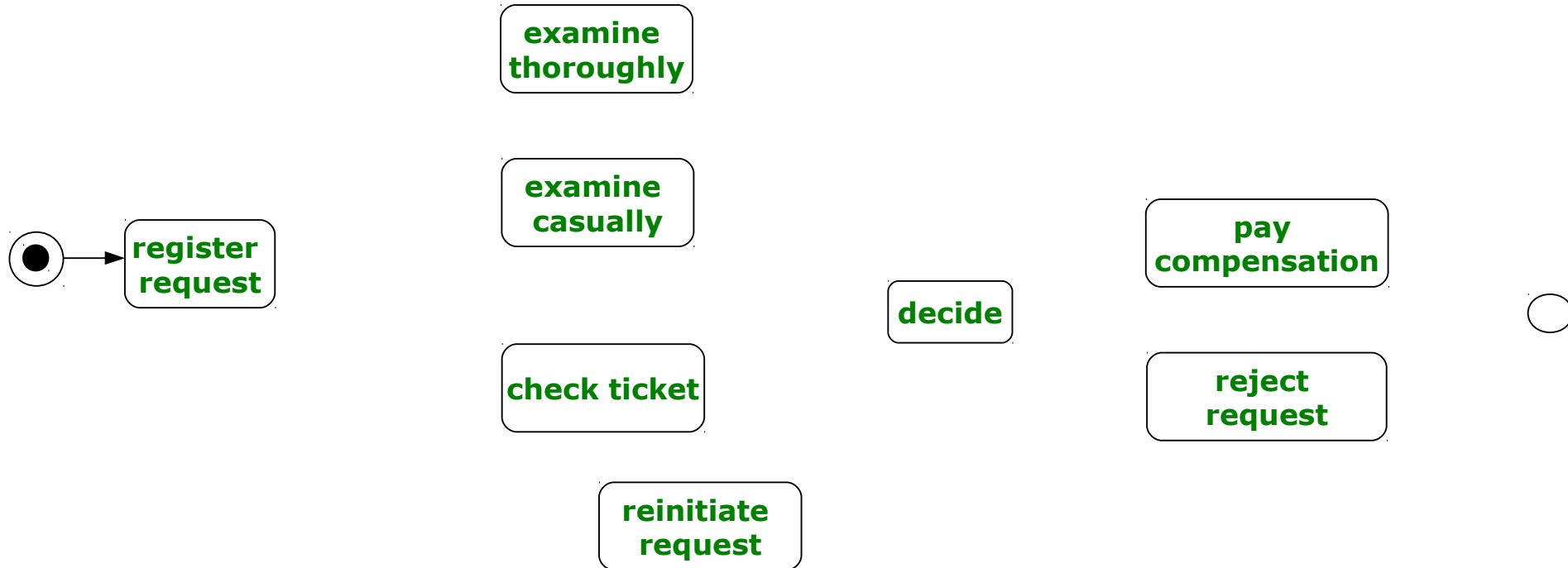
# Example Process



- The process shows only one possible workflow network

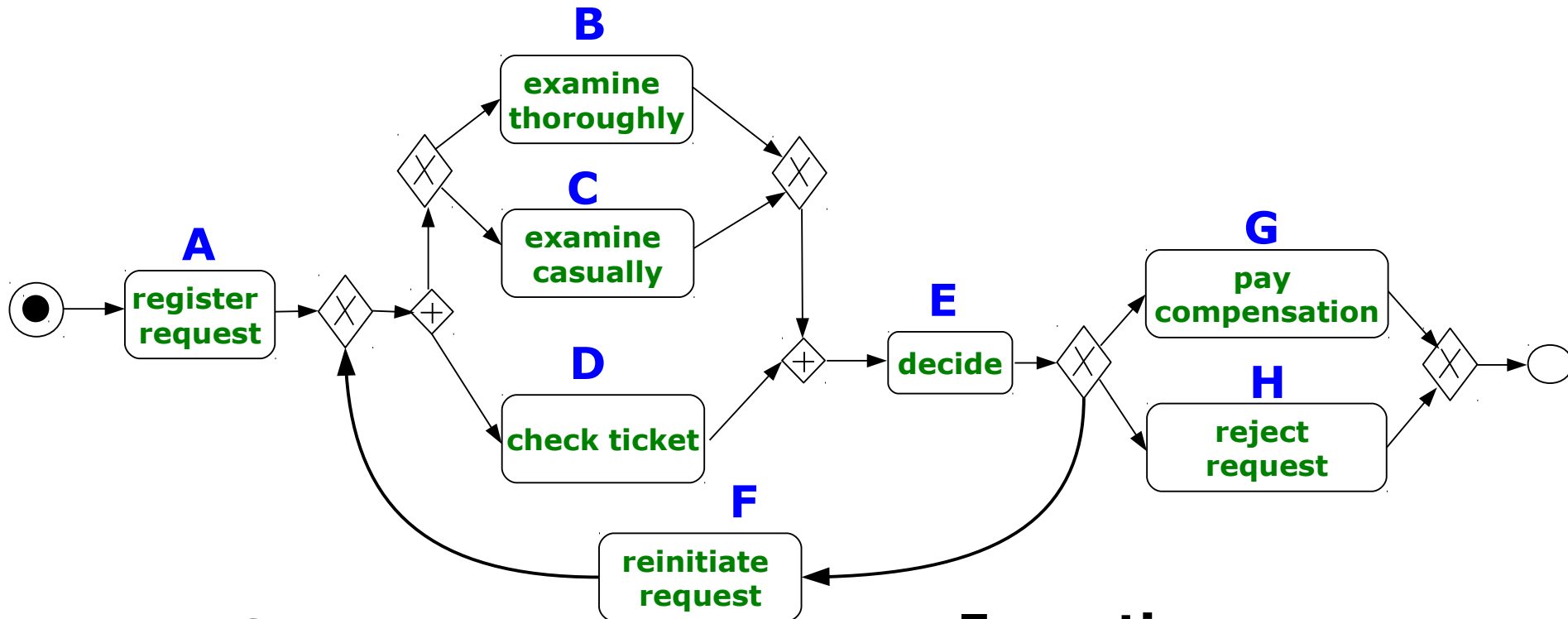
Example adapted from  
(van der Aalst et al., 2011)

# Activities without the Workflow Network



Users are not **forced** to follow the process

# Example Activity Logs



## • Event Sequence:

- {a, b, e, g}
- {a,c,e,h}
- {a,d,e,g,c,h}
- {a,b,e,f,d,e,f,d,e,g}

## • Exceptions

- {a, e, g}
- {a, e, h}
- {a, g}
- {a,e,h}

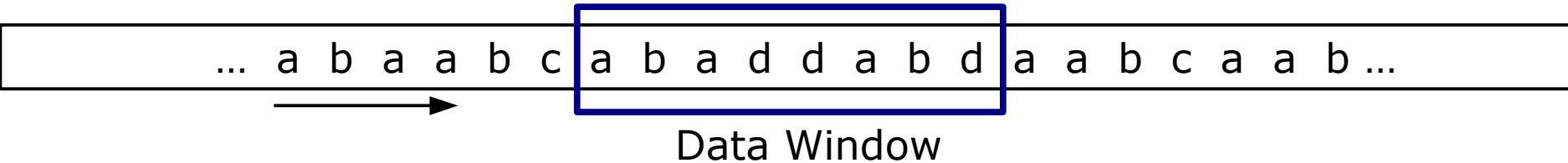


# Patterns of Complex Events

Event Types: A, B, C, D

Event Instances: a, b, c, d

Example Event Stream:



Detection of **Complex Events** based on **predefined Patterns**

Pattern (A **AND** B **AND** D) : 5 Times

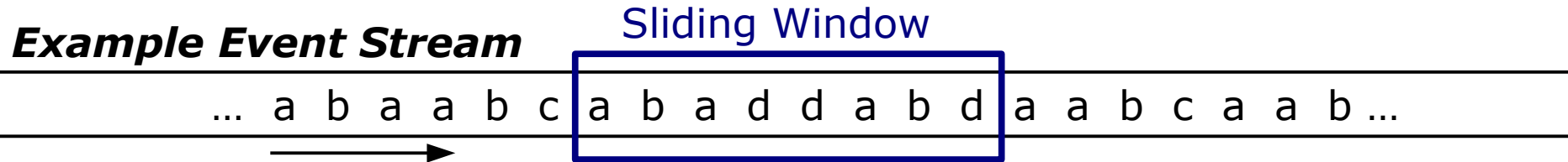
Pattern (A **OR** B **OR** D) : 9 Times

Pattern (A **SEQ** B **SEQ** D) : 3 Times

Pattern (**NOT** C) : 1 Time

# Event Pattern Mining

Types: A, B, C, D  
 Instances: a, b, c, d



- We do not have any patterns.
- We are interested in finding patterns from data.
  - Anomaly Detection on Event Stream
  - Looking for Event Patterns based on Event Operators
  - Frequent Patterns

Which Patterns are interesting here?

**Pattern: ( A SEQ D SEQ D SEQ A )**

# Complex Event Pattern

- Examples of **Complex Events**
  - abnormal activity in an on-line banking website
  - abnormal credit approval in a bank
  - abnormal flight ticket compensation
  
- Complex Event Pattern
  - based on **sequence of activities**
    - event algebra operators like SEQ, AND, OR,...
  
  - based on **relations of activities** to other resources
    - using background knowledge about actors, activities, events, ...

# Metadata about Process Activities

**A**

register request

- Ticket id
- Actor
- Time/Date
- Event (strikes, bad wetter)
- Reregistration

**B**

examine thoroughly

- Ticket
- Actor
- Time/Date
- Reasons
- ...

**C**

examine casually

**D**

check ticket

- From/To
- Actor
- Time/Date
- ...

decide **E**

**F**

reinitiate request

pay compensation **G**

reject request **H**

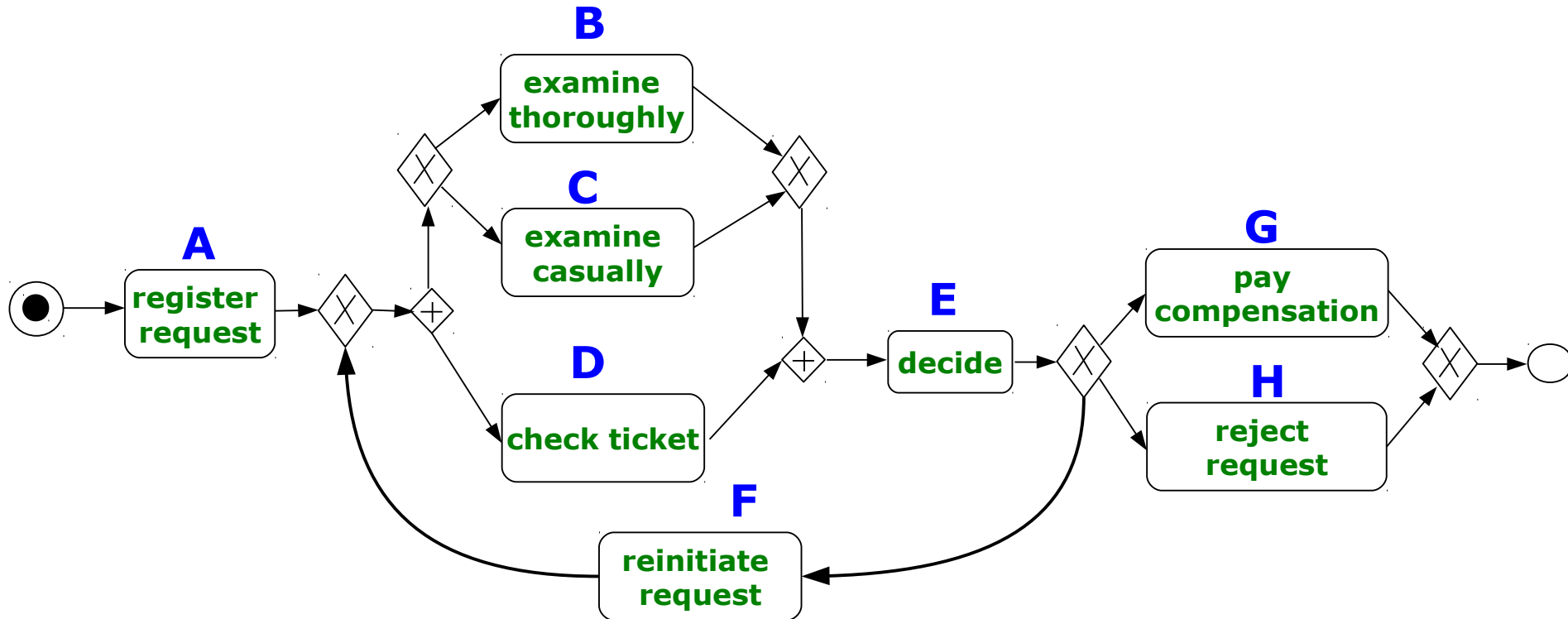
## General Attributes

- Time
- Actors

## Ticket Attributes

- Passenger (e.g., VIPs)
- Customer (e.g., special)
- Destination
- Fare Type  
(economy, business, first)

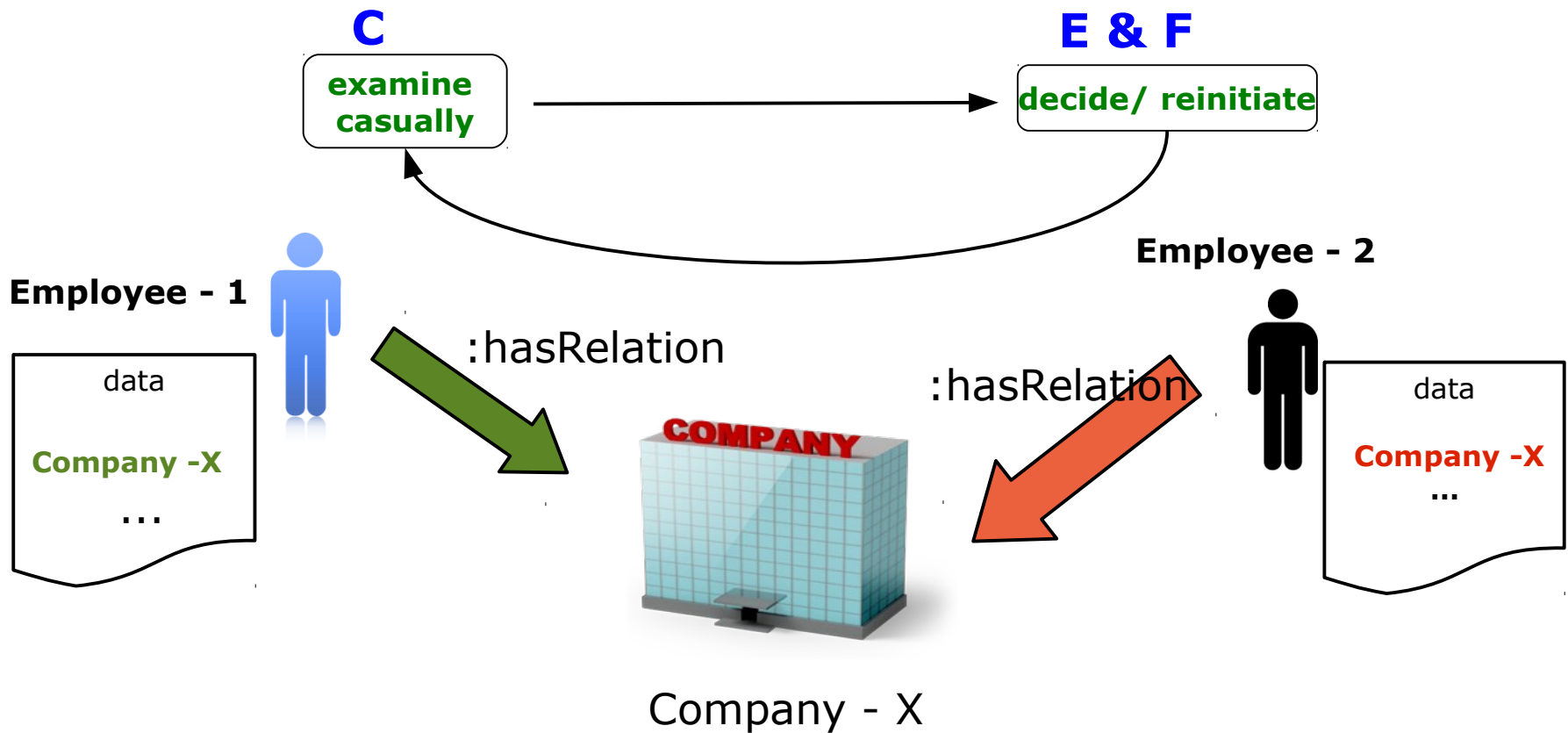
# Example Activity Logs – Loop Sequence



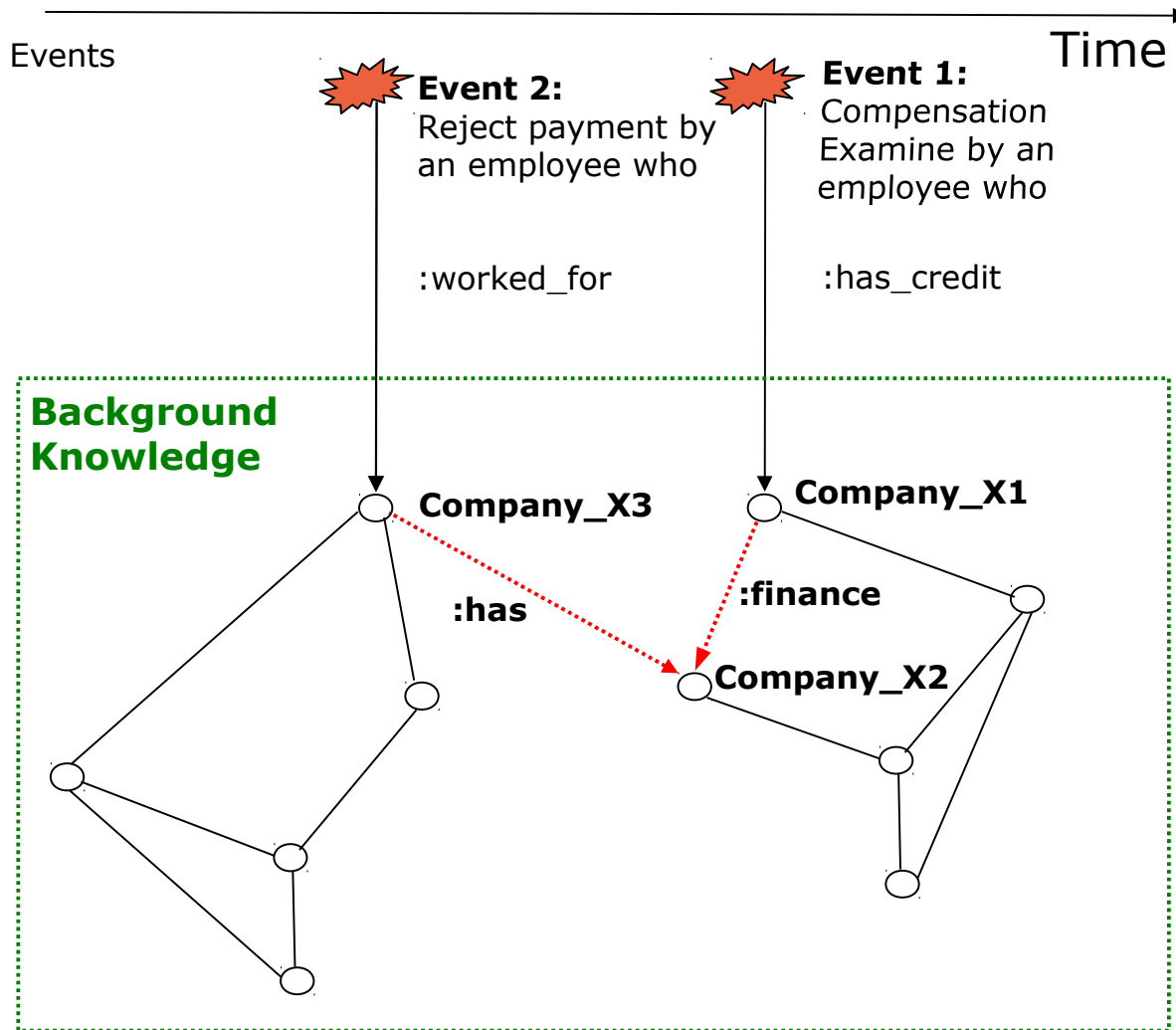
- Event Sequence {a,c,d,e,f,c,d,e,f,b,d,e,g}

# Complex Event Pattern

- A sequence of ticket compensation proof and rejection



# Patterns Based on Background Knowledge



Pattern:

**[Register Request/examine]**  
**→ SEQ**  
**[Payment Decision]**

- Examine by an employee who worked for a company. That Employee is fired by the company X2.

- Compensation accept by an employee who has credit from a company X1

# Semantic Sequential Pattern Mining

- Mining of complex event patterns based on:
  - Temporal relations of events
  - Relations of events to other sources like actors
  
- Ontological reasoning on
  - actions, actors, events, situations, ...
  
- Enabling extraction of complex high-level business events
  
- Discovery of Workflow driven by complex events



# Research Questions

- **Enrichment of Event Logs**
  - To which extent should we extract knowledge about events?
  - Complete Enrichment vs. Partial Enrichment
  - Partial enrichment for special kind of graph patterns
- **Integrate the Sequential Pattern Mining**
  - Mining sequence of *knowledge graphs patterns*
  - Integration ontological reasoning in sequential pattern detection algorithms
- **Real-Time Pattern Detection**
  - Algorithms for high throughput *knowledge graph streams*
  - Usage of sketching and sampling algorithms

# Conclusion

- Usage of background knowledge for abstract pattern specification
- Enabling knowledge-based anomaly detection from activities logs
- Enabling knowledge-based Process Discovering

# Thank You!

## Questions?



### **Acknowledgements**

This work has been partially supported by the “**InnoProfile-Transfer Corporate Smart Content**” project funded by the German Federal Ministry of Education and Research (**BMBF**) and the **BMBF Innovation Initiative** for the New German Länder “Entrepreneurial Regions”.