

Advanced Process Mining

Prof. Dr. Agnes Koschmider

Lecture 8: Applications



Lecture Overview

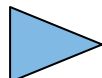
0 Organization and Introduction

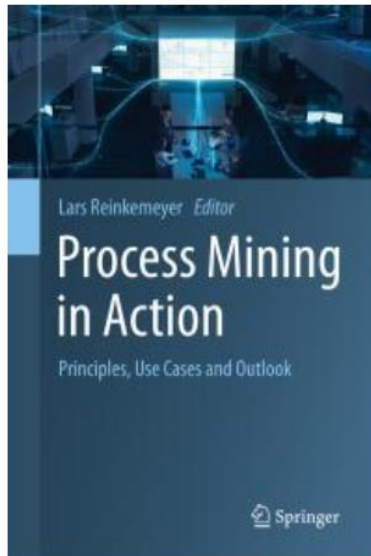
I Process Discovery

II Process Conformance

III Predictive Process Monitoring

IV Event Log Preparation

 V Applications and Practical Tasks

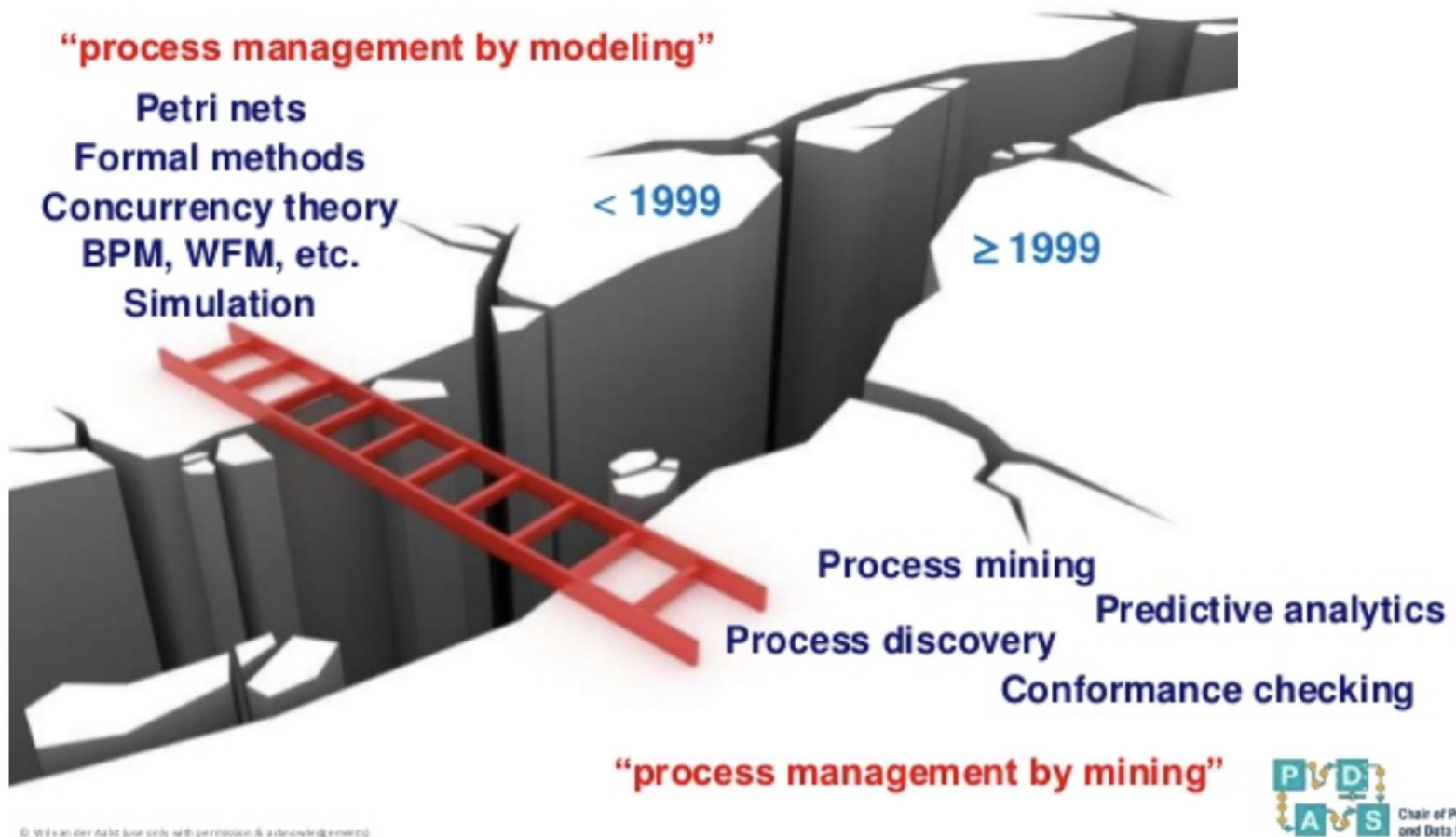


Process Mining in Action : Principles, Use Cases and Outlook

Lars Reinkemeyer

<https://ebookcentral.proquest.com/lib/christianalbrechts/detail.action?docID=6134217>

Process Management by modeling/mining



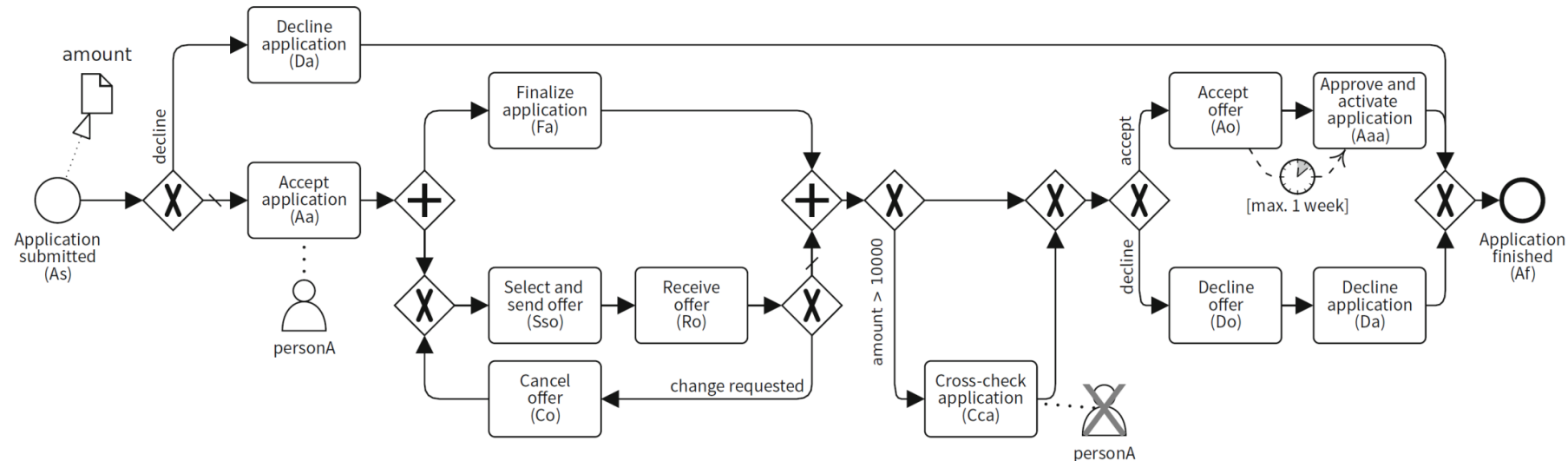
First and foremost, processes define behaviour, it is primarily about the ordering of activities

Yet, additional perspectives may also be relevant:

- Data: Which data objects are produced or consumed by an activity or checked at a decision point?
- Resources: Which people or systems are involved in the execution of an activity?
- Time: What temporal constraints affect the execution of activities?

These perspective may be important for conformance checking

Process Perspectives in a Model



Loan application processing

- Application has an amount
- Cross-checking only required for amounts > 10k
- Four-eyes-principle: Acceptance and cross-checking cannot be done by the same person
- At most a week between offer acceptance and activation of application

Traces that are conformant according to the control-flow may be non-conforming in according to other perspectives:

- Decision making based on data is not in line with model
- Resource assignments violate constraints imposed by model
- Temporal requirements are not satisfied

Considering further perspectives enables a more fine-granular view on conformance

Question of how to balance between the different perspectives

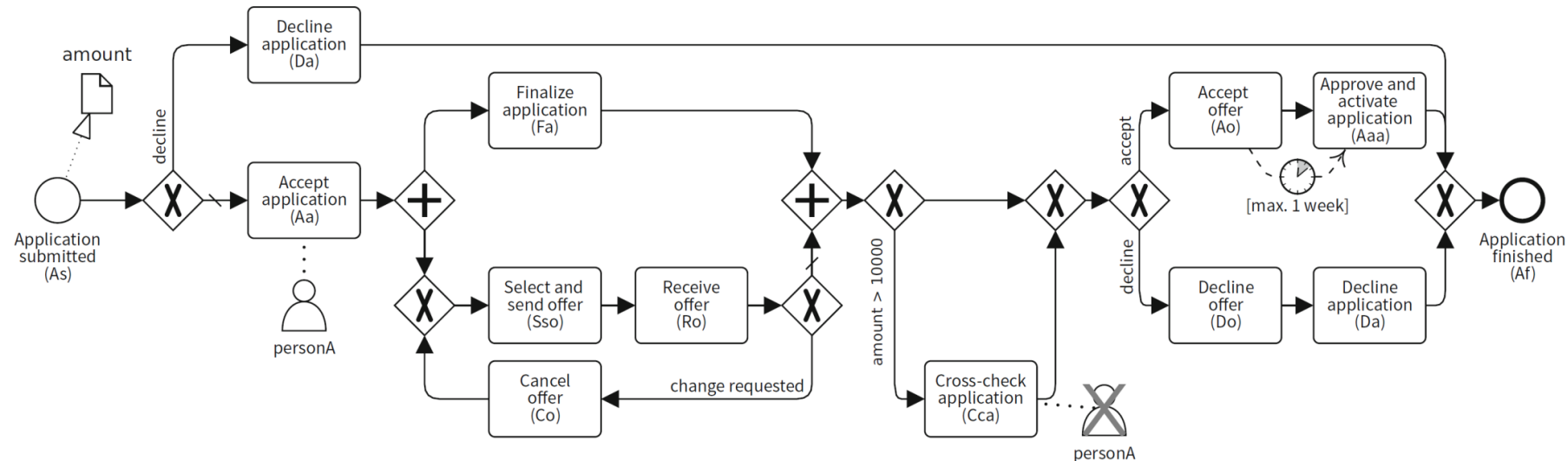
Example Continued(1)

Event	Application	Offer	Activity	Amount	Signed	Resource	Timestamp
<i>e</i> ₁₃	A0054		Application submitted	€12,000		<user>	Feb. 1, 12:31
<i>e</i> ₁₄	A0054		Accept application			John	Feb. 1, 12:32
<i>e</i> ₂₂	A0054		Finalise application			John	Feb. 3, 09:00
<i>e</i> ₃₇	A0054	5401	Select and send offer	€11,000		John	Feb. 5, 12:32
<i>e</i> ₄₂	A0054	5401	Receive offer		YES	John	Feb. 10, 10:00
<i>e</i> ₅₄	A0054	5401	Accept offer			John	Feb. 10, 10:04
<i>e</i> ₆₄	A0054		Approve and activate application			John	Feb. 18, 09:05
<i>e</i> ₆₅	A0054		Application finished				Feb. 18, 09:06

Using the abbreviations given in the model, the trace when looking only at the control-flow is:

$\langle As, Aa, Fa, Sso, Ro, Ao, Aaa, Af \rangle$

Example Continued(2)



$\langle As, Aa, Fa, Sso, Ro, Ao, Aaa, Af \rangle$

But:

- Amount: 11k => Cross-check application should have been executed
- Time between acceptance of offer and approval and activation of application nearly 8 days => should have been at most a week

Idea: Lift alignments from control-flow to multiple process perspectives

- “Step” in the alignment relates attribute values of a trace of an event log and of an execution sequence of a model to each other
- Attributes are: Activity, data, resources, time, ...
- Assigning costs to steps, optimal alignments are defined as before

log trace	execution sequence
<i>As</i> @Feb. 1, 12:31 (Res: <user>, Amount: €12,000)	<i>As</i> @Feb. 1, 12:31 (Res: <user>, Amount: €12,000)
<i>Aa</i> @Feb. 1, 12:32 (Res: John)	<i>Aa</i> @Feb. 1, 12:32 (Res: John)
<i>Fa</i> @Feb. 3, 09:00 (Res: John)	<i>Fa</i> @Feb. 3, 09:00 (Res: John)

Different perspectives provide different explanations for conformance issues

Consider the example of a 11k loan application for which cross-checking has not been done

- The amount may have been incorrectly recorded, it was actually 9k only, so that no cross-checking is fine
- The amount is indeed 11k, and the issue is the lack of cross-checking

However, costs assigned to the alignment of attribute values may prefer certain explanations (i.e., steps in the alignment)

Alignments for the Example

Alignment a_1	
log trace	execution sequence
<i>As</i> @Feb. 1, 12:31 (Res: <user>, Amount: €12,000)	<i>As</i> @Feb. 1, 12:31 (Res: <user>, Amount: €12,000)
<i>Aa</i> @Feb. 1, 12:32 (Res: John)	<i>Aa</i> @Feb. 1, 12:32 (Res: John)
<i>Fa</i> @Feb. 3, 09:00 (Res: John)	<i>Fa</i> @Feb. 3, 09:00 (Res: John)
<i>Sso</i> @Feb. 5, 12:32 (Amount: €11,000)	<i>Sso</i> @Feb. 5, 12:32 (Amount: € 10,000)
<i>Ro</i> @Feb. 10, 10:00 (Signed: YES)	<i>Ro</i> @Feb. 10, 10:00 (Signed: YES)
<i>Ao</i> @Feb. 10, 10:04 (Res: John)	<i>Ao</i> @Feb. 10, 10:04 (Res: John)
<i>Aaa</i> @Feb. 11, 09:05 (Res: John)	<i>Aaa</i> @Feb. 10, 11:03 (Res: John)
<i>Af</i> @Feb. 11, 09:06	<i>Af</i> @Feb. 18, 09:06

Alignment a_2	
log trace	execution sequence
<i>As</i> @Feb. 1, 12:31 (Res: <user>, Amount: €12,000)	<i>As</i> @Feb. 1, 12:31 (Res: <user>, Amount: €12,000)
<i>Aa</i> @Feb. 1, 12:32 (Res: John)	<i>Aa</i> @Feb. 1, 12:32 (Res: John)
<i>Fa</i> @Feb. 3, 09:00 (Res: John)	<i>Fa</i> @Feb. 3, 09:00 (Res: John)
<i>Sso</i> @Feb. 5, 12:32 (Amount: €11,000)	<i>Sso</i> @Feb. 5, 12:32 (Amount: €11,000)
<i>Ro</i> @Feb. 10, 10:00 (Signed: YES)	<i>Ro</i> @Feb. 10, 10:00 (Signed: YES)
»»	<i>Cca</i> @Feb. 10, 10:01 (Res: Carla)
<i>Ao</i> @Feb. 10, 10:04 (Res: John)	<i>Ao</i> @Feb. 10, 10:04 (Res: John)
<i>Aaa</i> @Feb. 11, 09:05 (Res: John)	<i>Aaa</i> @Feb. 10, 11:03 (Res: John)
<i>Af</i> @Feb. 11, 09:06	<i>Af</i> @Feb. 18, 09:06

Example Cost Function

Extend standard cost function by penalizing:

- Model moves that write to variables, where the number of written values adds to the cost,
- Synchronous moves that refer to different variables, where the number of different written variables adds to the cost

Revisit Example Alignments(1)

The amount was reduced to allow skipping the required cross-check activity. (+1)

The timestamp of the approve and activate application activity was adjusted to fit to the time constraint of less than one week. (+1)

Alignment a_1	
log trace	execution sequence
<i>As</i> @Feb. 1, 12:31 (Res: <user>, Amount: €12,000)	<i>As</i> @Feb. 1, 12:31 (Res: <user>, Amount: €12,000)
<i>Aa</i> @Feb. 1, 12:32 (Res: John)	<i>Aa</i> @Feb. 1, 12:32 (Res: John)
<i>Fa</i> @Feb. 3, 09:00 (Res: John)	<i>Fa</i> @Feb. 3, 09:00 (Res: John)
<i>Sso</i> @Feb. 5, 12:32 (Amount: €11,000)	<i>Sso</i> @Feb. 5, 12:32 (Amount: € 10,000)
<i>Ro</i> @Feb. 10, 10:00 (Signed: YES)	<i>Ro</i> @Feb. 10, 10:00 (Signed: YES)
<i>Ao</i> @Feb. 10, 10:04 (Res: John)	<i>Ao</i> @Feb. 10, 10:04 (Res: John)
<i>Aaa</i> @Feb. 11, 09:05 (Res: John)	<i>Aaa</i> @Feb. 10, 11:03 (Res: John)
<i>Af</i> @Feb. 11, 09:06	<i>Af</i> @Feb. 18, 09:06

Revisit Example Alignments(2)

A model move for the mandatory activity to cross-check the application was added, to reflect the four-eyes-principle before accepting the offer. (+1)

- A time value was added to the model move. (+1)
- A resource value was added to the model move. (+1)

The timestamp of the activity to approve and activate the application was adjusted to fit to the time constraint of less than one week. (+1)

Alignment a_2	
log trace	execution sequence
<i>As @Feb. 1, 12:31</i> (Res: <user>, Amount: €12,000)	<i>As @Feb. 1, 12:31</i> (Res: <user>, Amount: €12,000)
<i>Aa @Feb. 1, 12:32</i> (Res: John)	<i>Aa @Feb. 1, 12:32</i> (Res: John)
<i>Fa @Feb. 3, 09:00</i> (Res: John)	<i>Fa @Feb. 3, 09:00</i> (Res: John)
<i>Sso @Feb. 5, 12:32</i> (Amount: €11,000)	<i>Sso @Feb. 5, 12:32</i> (Amount: €11,000)
<i>Ro @Feb. 10, 10:00</i> (Signed: YES)	<i>Ro @Feb. 10, 10:00</i> (Signed: YES)
»	Cca @Feb. 10, 10:01 (Res: Carla)
<i>Ao @Feb. 10, 10:04</i> (Res: John)	<i>Ao @Feb. 10, 10:04</i> (Res: John)
<i>Aaa @Feb. 11, 09:05</i> (Res: John)	Aaa @Feb. 10, 11:03 (Res: John)
<i>Af @Feb. 11, 09:06</i>	<i>Af @Feb. 18, 09:06</i>

Alignments for Event Log Analysis

So far, alignments between a trace of an event log and an execution sequence of a model

- Pairwise alignment
- Each trace is considered in isolation

But: Alignments are also useful in the absence of a model

- Support explorative analysis of an event log
- Insights into regularities and most common behaviour among the traces
- Identification of rare deviations
- Powerful visual feedback

Construct alignment between traces of event log

- No longer a question of optimising the alignment cost for a single trace
- Global view: overall alignment cost should be minimal

Problem well-known in genomics

- Alignments of nucleic acid sequences
- Yet, also known to be an NP-complete problem

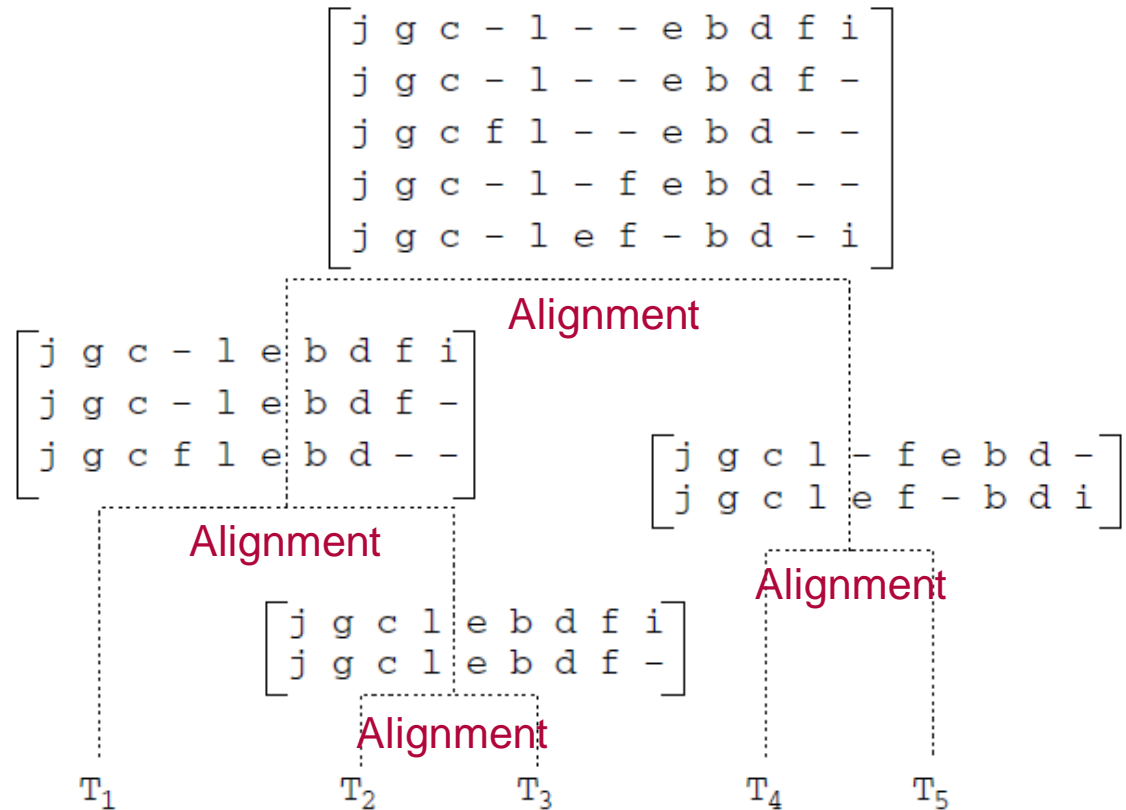
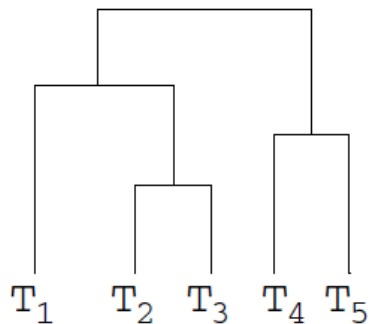
Various heuristic techniques to find multiple trace alignment that may be non-optimal

- Typically based on iterative approach
- Often based on hierarchical clustering

Intuition

T_1 : j g c f l e b d
 T_2 : j g c l e b d f i
 T_3 : j g c l e b d f
 T_4 : j g c l f e b d
 T_5 : j g c l e f b d i

Hierarchical
Clustering



Standard techniques may lead to results that are not intuitive from a business process point of view

Example: tendency to penalise gaps at the beginning and end of traces

- Not in line with “looping” behaviour in processes
- Consider locality in the alignment

j g c - a h b - - - - f d

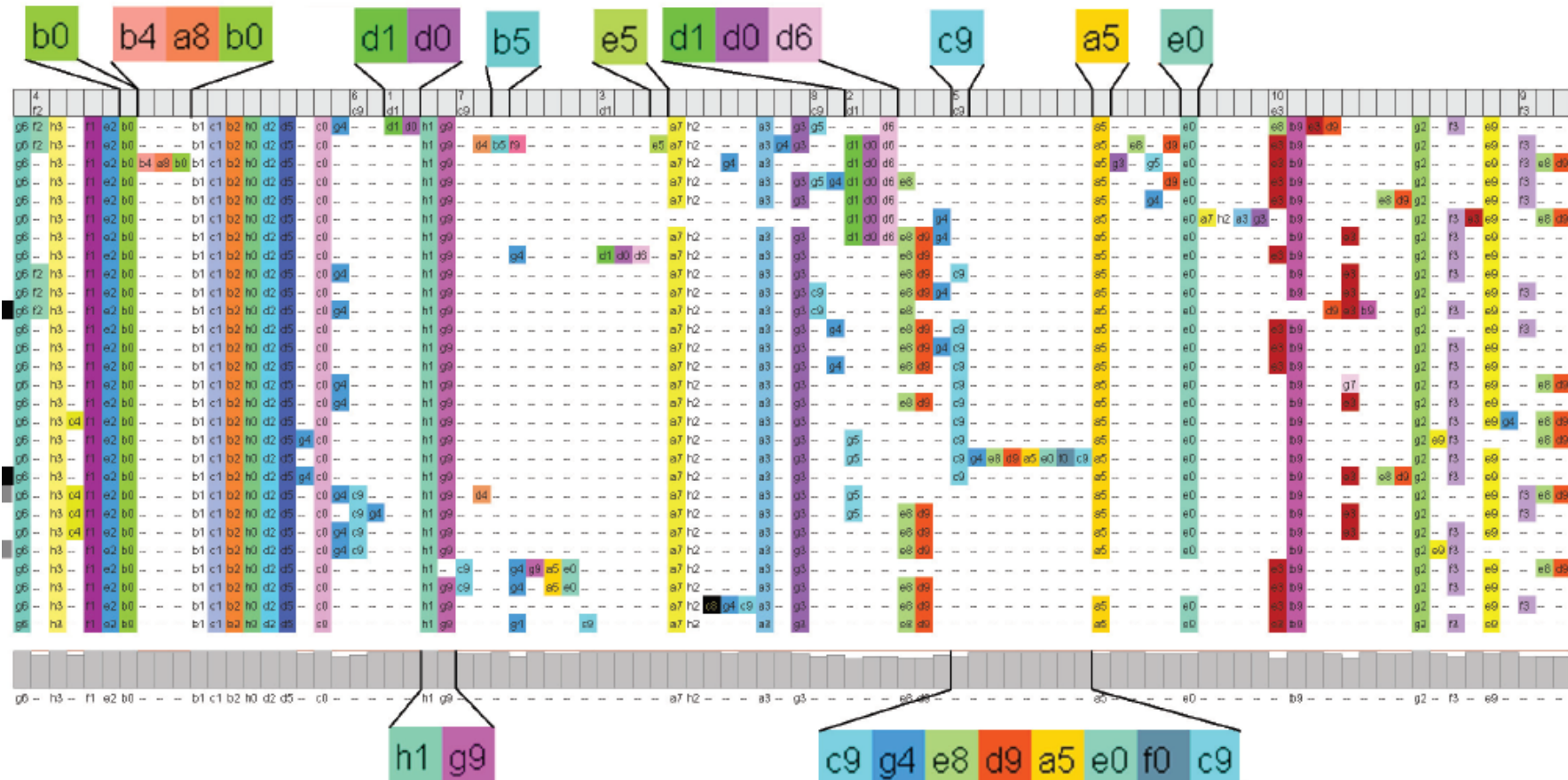
j g c f a h b d k a h b d

j g c - a h b f d - - - - -

j g c f a h b - d k a h b d

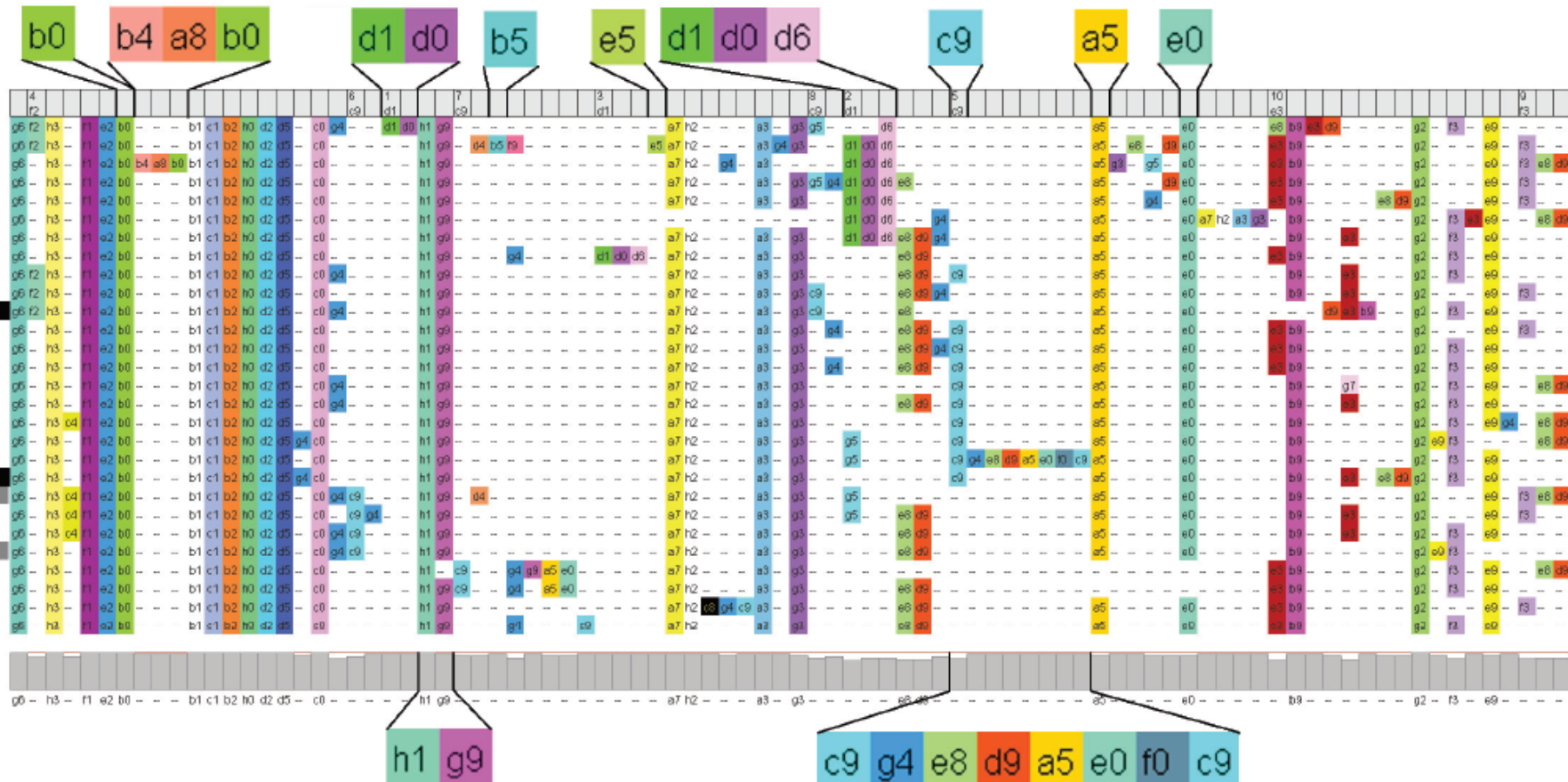
Insights(1)

Alignment of a cluster of real world traces (house/apartment rentals):



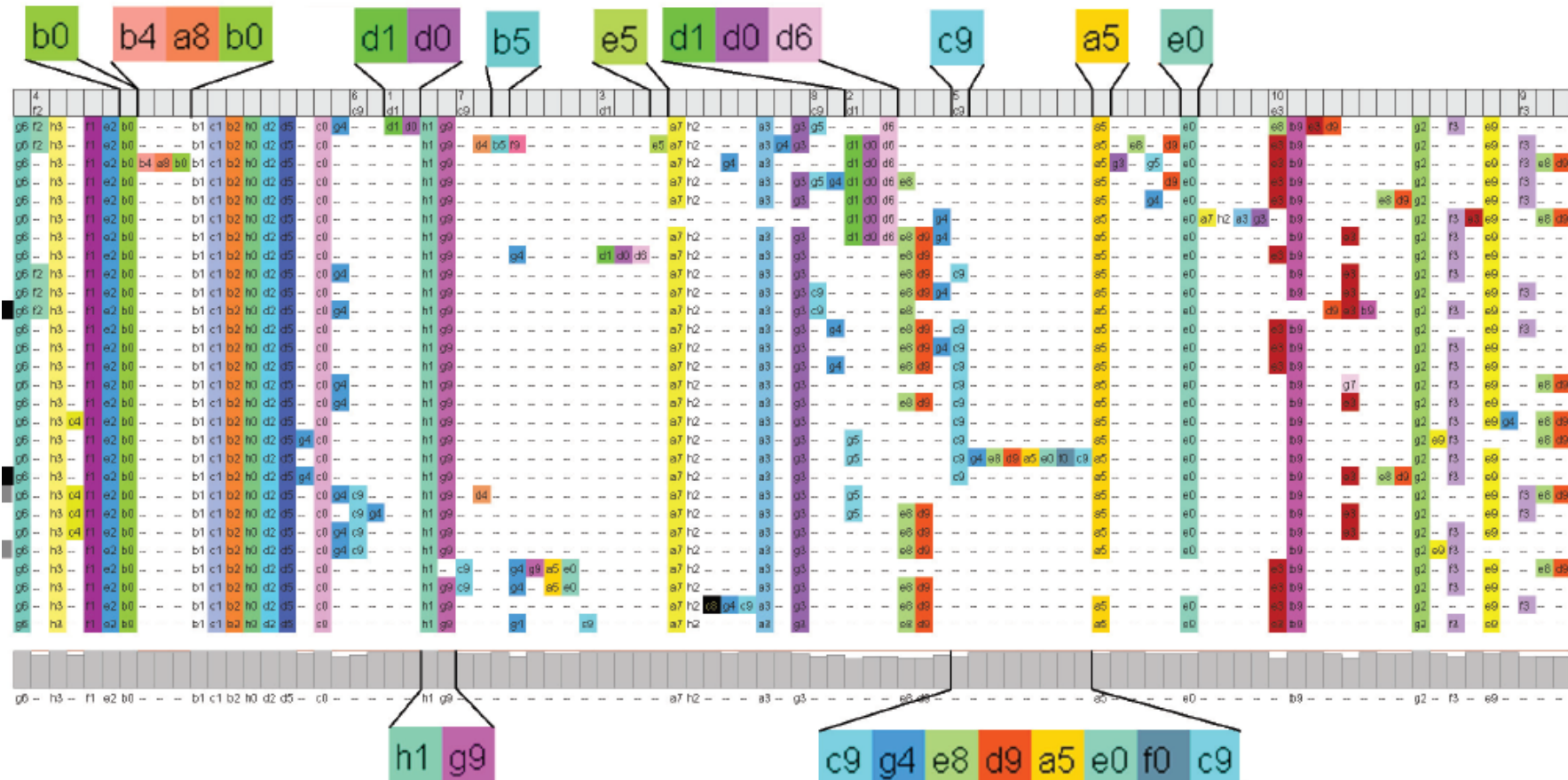
Insights(2)

b0 is check activity (inspection done?), sequence b4a8b0 is done only in negative case (tenant not at home -> send letter, reschedule)



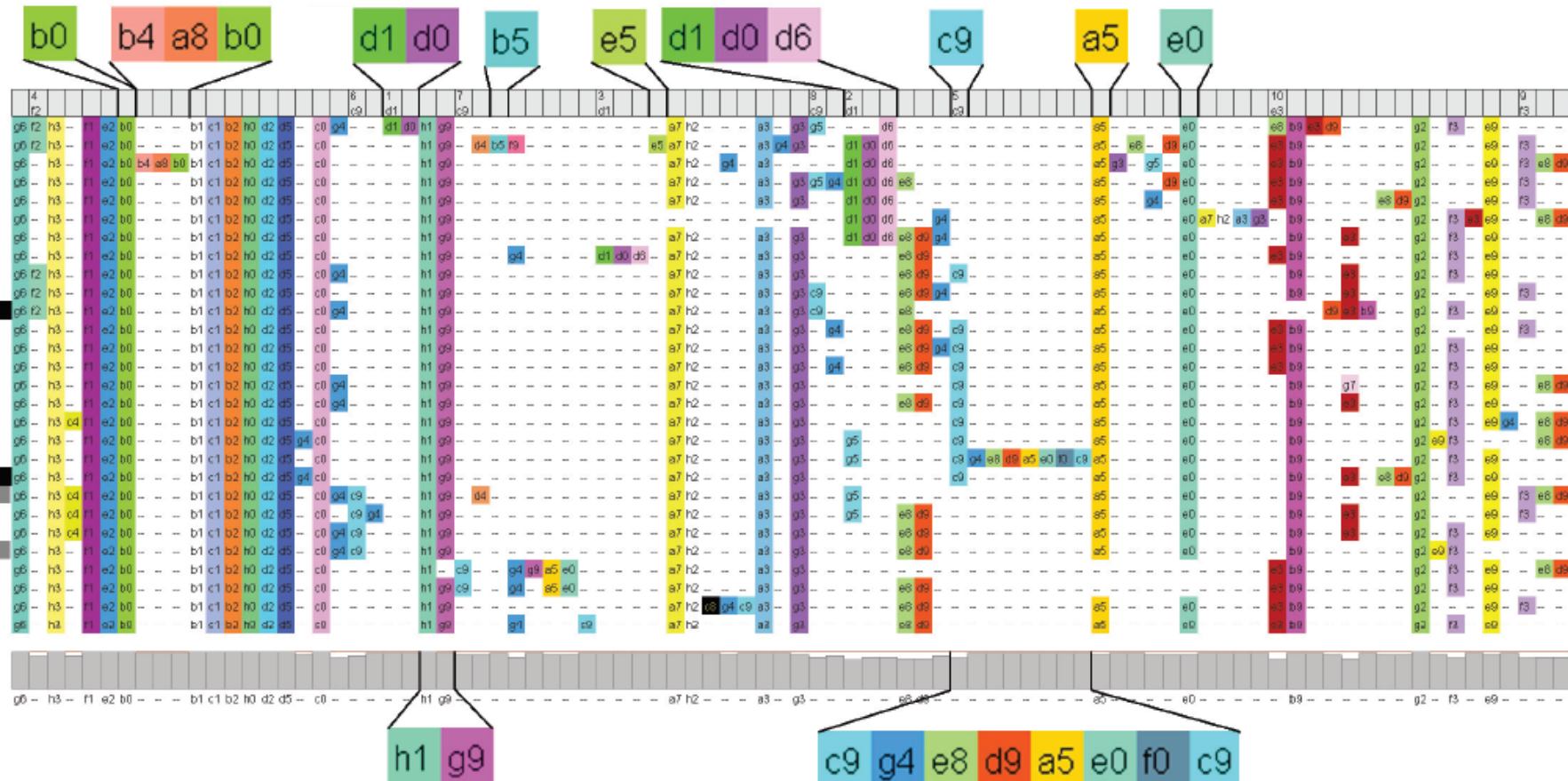
Insights(3)

h1g9 is well preserved (final inspection, recording of defects)



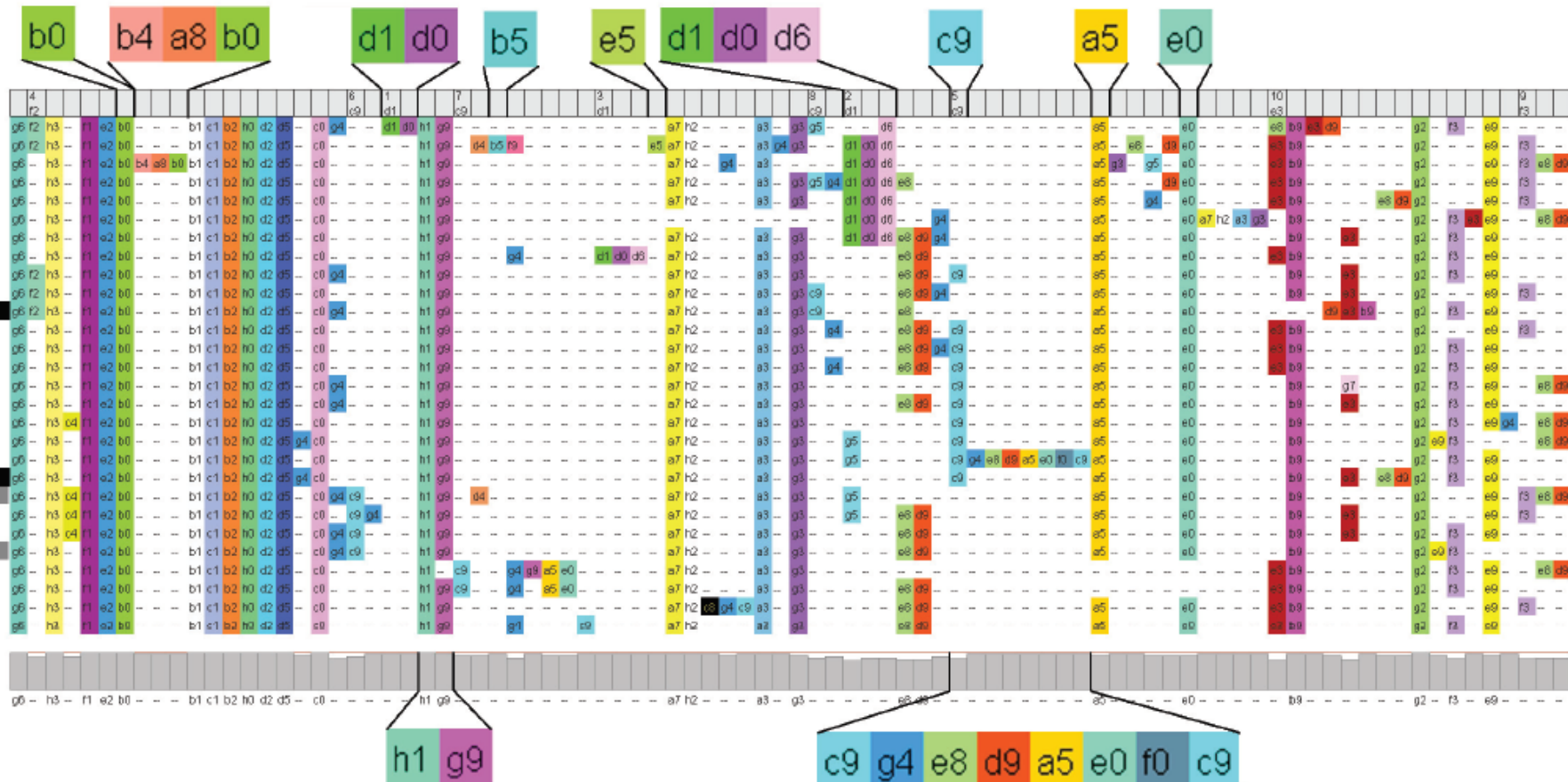
Insights(4)

d1d0 is offering a flat, before final inspection in first trace!

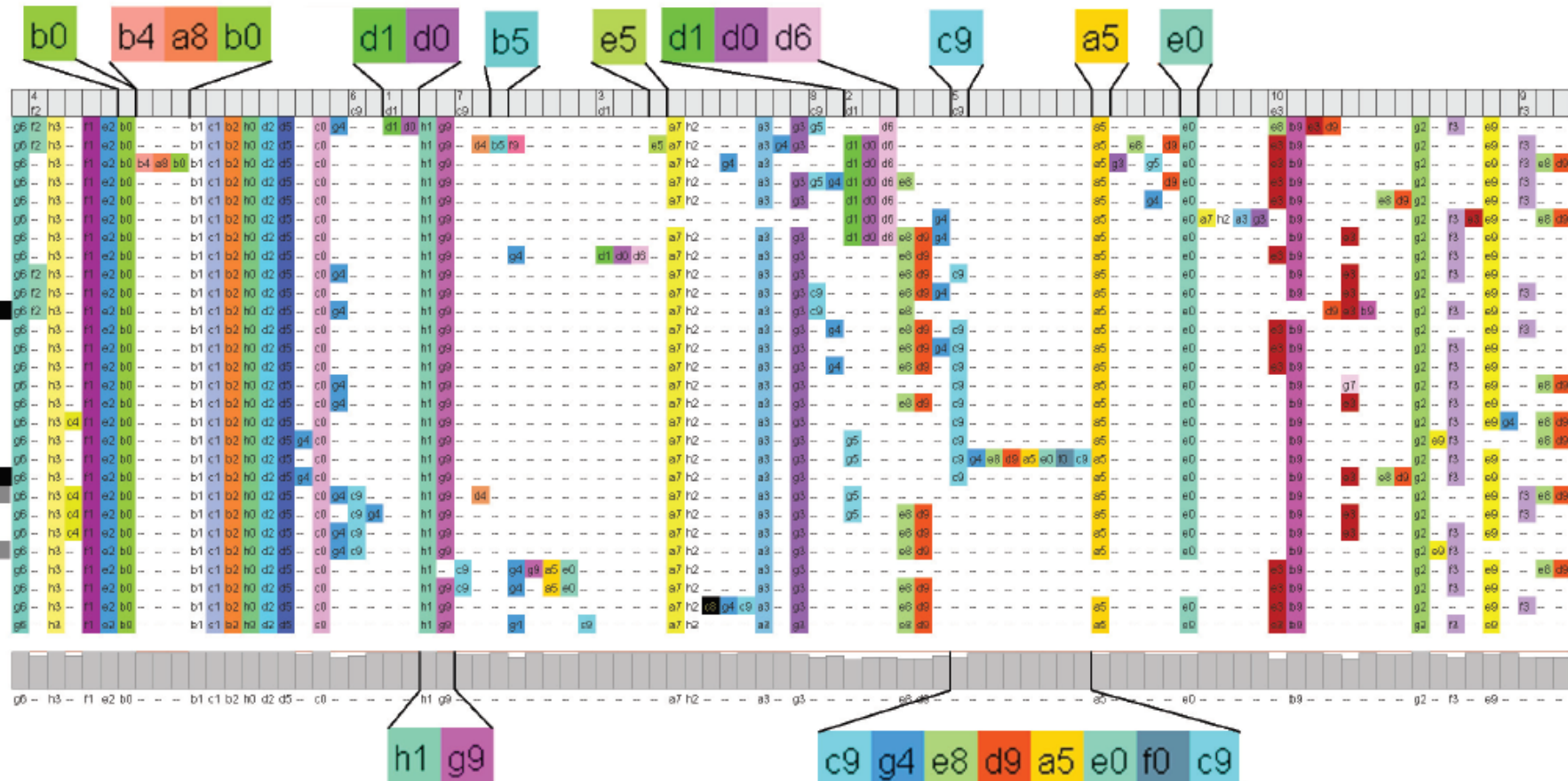


Insights(5)

b5e5 are related to second inspection – rarely needed.



c9 is determination of candidate tenant – happened multiple times in one trace (featuring f0 termination of provisional lease)



Identifying the Right Use Cases

- quality of the data presented to process modeling algorithms is critical to the success of any process mining exercise
- Pre-processing (cleaning) event logs to address quality issues prior to conducting a process mining analysis is necessary, but time-consuming task

Key learnings from industry settings(1)

#1 Transparency is a prerequisite for digital transformation:

Digital transformation should be unbiased, transparent, and measurable.

Process Mining allows transparency on actual processes and provides insights for fact-based digital transformation.

Key learnings from industry settings(2)

#2 Process Mining allows full transparency based on event logs:

An event log is a collection of events which have taken place in order to perform a business process.

Event logs allow to visualize process variants in the actual sequence of activities.

Key learnings from industry settings(2)

#3 Purpose, People, and Processes are essential (“3Ps”):

All three factors have to be in sync for successful projects and can't substitute each other.

Purpose: Identifying the Right Use Cases

- a clear understanding is needed of what Process Mining shall be used for
- Purpose may be defined to achieve higher customer satisfaction, less rework, faster deliveries, lower transactional cost, less traffic, or less waste
- Purpose comes first and needs to be clearly articulated in the form of use cases

- it is a key success factor to get the right people on board
- process experts or process owners have to appreciate and support the approach
- The process starts with the right process owner who is passionate about change and innovation
- encounter experts who have managed a process for many years and who are reluctant to change
- Avoid misalignment between purpose and approach, otherwise surrogation may occur
- a close collaboration between functional departments and IT is recommended

- the identification of relevant digital traces and customization of reliable data models is required
- standardize and harmonize the ingested raw data
- the normalized event logs are connected to actual process flows and visualized in frontend reports
- Provide an Open Platform to Build a Strong Community

Key learnings from industry settings

#4 Start with simplicity to fight complexity:

Clear and simple project targets should be defined to allow for a successful start.

After a successful start, the scope can be extended to achieve transparency on process complexity.

Key learnings from industry settings

#5 Purpose comes first:

At the beginning of any project a purpose with a clear business value must be defined, typically in the form of a use case.

The purpose must be achievable and supported by the process owner and all involved team members.

Key learnings from industry settings

#6 It's all about the people:

Engaging the right passionate drivers in operational units, who understand how Process Mining can improve operational efficiency and drive digitalization, is crucial.

Key learnings from industry settings

#7 Processtraces are comparable to raw oil:

You have to search for it, find it, collect it, and refine it.
In the right quality, it can become a very powerful fuel.

Key learnings from industry settings

#8 Provide an open platform and build a strong community:

Process Mining services should be accessible to every employee, with an open exchange of experiences.

Key learnings from industry settings

#9 Fail fast or scale fast:

An agile project setup should allow quick results and checkpoints.

Fast scaling of services should be considered in every project phase.