Guided Interaction Exploration in Artifact-centric Processes

M.L. van Eck,
N. Sidorova,
W.M.P. van der Aalst
Process Discovery

Event Log → Process Discovery Algorithm → Process Model
Process Discovery
Artifact-centric Processes
Example Process

Process of Life

Asleep, not hungry → Awake, hungry

Awake, not hungry ↔ Awake, eating

Nutrition

Hungry
→ Eating
→ Not hungry

Sleep

Asleep
→ Awake

Example Process

Nutrition

Hungry
→ Eating
→ Not hungry

Sleep

Asleep
→ Awake
### Process Instances

<table>
<thead>
<tr>
<th>Instance</th>
<th>Start</th>
<th>End</th>
<th>State</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>6:00</td>
<td>7:00</td>
<td>Asleep, Not hungry</td>
</tr>
<tr>
<td>1</td>
<td>7:00</td>
<td>8:00</td>
<td>Awake, Hungry</td>
</tr>
<tr>
<td>1</td>
<td>8:00</td>
<td>8:30</td>
<td>Awake, Eating</td>
</tr>
<tr>
<td>1</td>
<td>8:30</td>
<td>12:00</td>
<td>Awake, Not hungry</td>
</tr>
<tr>
<td>1</td>
<td>12:00</td>
<td>12:30</td>
<td>Awake, Eating</td>
</tr>
<tr>
<td>1</td>
<td>12:30</td>
<td>17:00</td>
<td>Awake, Not hungry</td>
</tr>
<tr>
<td>1</td>
<td>17:00</td>
<td>18:00</td>
<td>Awake, Hungry</td>
</tr>
<tr>
<td>1</td>
<td>18:00</td>
<td>19:00</td>
<td>Awake, Eating</td>
</tr>
<tr>
<td>1</td>
<td>19:00</td>
<td>23:00</td>
<td>Awake, Not hungry</td>
</tr>
<tr>
<td>1</td>
<td>23:00</td>
<td>0:00</td>
<td>Asleep, Not hungry</td>
</tr>
<tr>
<td>2</td>
<td>5:00</td>
<td>7:30</td>
<td>Asleep, Not hungry</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>
Guided Interaction Exploration
CSM Miner
Guided Interaction Exploration
Artifact Interaction

• **Examples:**
  - Causal dependencies between activities in different artifacts
  - Synchronization conditions
  - Co-occurrence of states
  - ...

Asleep, not hungry  ➔  Awake, hungry

Asleep  ➔  Awake, eating

Awake, eating  ➔  Eating  ➔  Not hungry  ➔  Awake

Hungry  ➔  Eating  ➔  Not hungry  ➔  Awake

Asleep  ➔  Awake, eating

Awake, not hungry  ➔  Hungry
Guided Interaction Exploration
Quantifying Artifact Interaction

- 3 types of interaction
- Implications over statements of artifact behaviour

Asleep, not hungry  Awake, hungry

Awake, not hungry  Awake, eating

Hungry 
Eating
Not hungry

Awake
Asleep
Quantifying Artifact Interaction

- State co-occurrence ($X \rightarrow Y$)
  - $X =$ artifact I is in state $s$
  - $Y =$ artifact J is in state $s'$

Asleep, not hungry  
Awake, hungry

Awake, not hungry  
Awake, eating

Hungry  
Eating

Not hungry  
Awake

Asleep
Quantifying Artifact Interaction

• State co-occurrence (X \(\Rightarrow\) Y)
  - X = artifact **Sleep** is in state **Awake**
  - Y = artifact **Nutrition** is in state **Eating**

• Quantify through probabilities
  - \(P(X) = \frac{\text{time spent Awake}}{\text{total time}}\)
  - \(P(Y) = \frac{\text{time spent Eating}}{\text{total time}}\)

• Confidence of X \(\Rightarrow\) Y:
  - \(P(Y|X) = \frac{\text{time spent Awake and Eating}}{\text{time spent Awake}}\)
Quantifying Artifact Interaction

- Transition co-occurrence ($X \rightarrow Y$)
  - $X =$ artifact I is changing state from $s$ to $s'$
  - $Y =$ artifact J is in state $s''$

Asleep, not hungry  
Awake, hungry  
Awake, not hungry  
Awake, eating  
Hungry  
Eating  
Not hungry  
Asleep  
Awake
Quantifying Artifact Interaction

• Transition co-occurrence (X ➔ Y)
  • X = artifact Sleep is changing state from Awake to Asleep
  • Y = artifact Nutrition is in state Eating

• Confidence of X ➔ Y:
  • \( P(Y|X) = \frac{\text{transition freq of (Awake and Eating) to (Asleep and Eating)}}{\text{transition freq of Awake to Asleep}} \)
Quantifying Artifact Interaction

- **Forward-looking co-occurrence (X \(\rightarrow\) Y)**
  - **X** = artifact I is in state \(s\) and artifact J is in state \(s'\)
  - **Y** = the next transition in artifact J is from state \(s'\) to \(s''\)

Asleep, not hungry  \(\rightarrow\)  Awake, hungry

Awake, not hungry  \(\rightarrow\)  Awake, eating

Hungry  \(\rightarrow\)  Eating

Eating  \(\rightarrow\)  Not hungry

Asleep  \(\rightarrow\)  Awake
Quantifying Artifact Interaction

- **Forward-looking co-occurrence** (X $\Rightarrow$ Y)
  - X = artifact **Sleep** is in state **Awake** and artifact **Nutrition** is in state **Not hungry**
  - Y = the next transition in artifact **Nutrition** is from state **Not hungry** to **Eating**

- **Confidence of X $\Rightarrow$ Y:**
  - \[ P(Y|X) = \frac{\text{transition freq of (Awake and Not hungry) to (Awake and Eating)}}{\text{total freq of Nutrition transitions from (Awake and Not hungry)}} \]
Measures of Interest

- **Objective measures of interestingness of correlations**
  - **Confidence** \( P(Y|X) \)
  - **Support** \( P(Y|X)P(X) = P(X \land Y) \)
  - **Lift** \( \frac{P(Y|X)}{P(Y)} \)
  - **Conviction** \( \frac{P(X)P(\bar{Y})}{P(X\land\bar{Y})} \)
  - **Cosine** \( \frac{P(X\land Y)}{\sqrt{P(X)P(Y)}} \)
  - **Jaccard** \( \frac{P(X\land Y)}{P(X\lor Y)} \)
  - **Phi-coefficient** \( \frac{P(X\land Y) - P(X)P(Y)}{\sqrt{P(X)P(Y)(1-P(X))(1-P(Y))}} \)
  - **...**

- **Many more from association rule learning**
Guided Interaction Exploration
Using Measures of Interest

- Survey to relate statistical measures and subjective interest

9.8% of the total time spent in Application Submitted (A) is spent while being in Process Leads (B) *
Using Measures of Interest

9.8% of the total time spent in Application Submitted (A) is spent while being in Process Leads (B)
Using Measures of Interest

Transitions from Application Finalized (A) to Application Cancelled (B) occur 90.4% of the times while being in Call After Offers (C)

Transitions from Application Finalized (A) to Application Approved (B) occur 91.3% of the times while being in Validation (C)
Using Measures of Interest

99.6% of the total time spent in Offer Cancelled (A) is spent while being in Application Cancelled (B)
Guided Interaction Exploration

• Demo with CSM Miner!
Conclusions

• Traditional discovery approaches have difficulties with complex artifact-centric processes

• Artifact-centric process mining gives simple models, but interaction between artifacts is still complex

• Guided exploration by ranking the relations

• Available as a plug-in for ProM 6.7: CSMMiner
Future Work

• Artifact interaction
  • Sets of artifacts instead of pairs
  • Currently only exact overlap in time
  • Dealing with gaps & limited data (e.g. partial data replay)

• Evaluation
  • Relation between objective measures and subjective interest
  • Determining thresholds for measures

• Multi-instance process mining
  • Artifacts with multiple case IDs
  • E.g. 1 order with 3 invoices
Questions?