DrFurby Classifier

Eric Verbeek
Felix Mannhardt

Where innovation starts
• Contest
• Observations
• Model
• Discovery
• Algorithm selection
• Extension
• Implementation
• Results
• Conclusions
• Contest
• Observations
• Model
• Discovery
• Algorithm selection
• Extension
• Implementation
• Results
• Conclusions
• 10 Models, not disclosed
• For every model:
  – March log containing 1000 positive traces
    • ‘Training’ log
  – April/May logs containing 10 positive and 10 negative traces
    • ‘Calibration’ logs
  – June log containing 10 positive and 10 negative traces
    • ‘Test’ log
• 200 classifications
  – Contribution with most correctly classified traces wins
  – In case of a tie, the fastest wins
• Contest
• Observations
• Model
• Discovery
• Algorithm selection
• Extension
• Implementation
• Results
• Conclusions
• The training logs contain no noise
  – All traces in the training logs are to be classified as positive
• The winner is the one that classifies the most traces correctly
  – The readability of the model is not relevant
• Contest
• Observations
• Model
• Discovery
• Algorithm selection
• Extension
• Implementation
• Results
• Conclusions
• A collection of discovered accepting Petri nets
  – Initial marking
  – Set of final markings
• Semantics:
  – Replay the trace-to-classify on every discovered net and accumulate the replay costs
    • Costs for move-on-log: 10
    • Costs for move-on-model: 4
    • Other costs: 0
  – Classify the trace-to-classify as positive if and only if the accumulated replay costs are 0
  – Use decomposed replay to speed up the replay
    • Decomposed replay preserves perfect fitness, that is, costs 0
• Contest
• Observations
• Model
• Discovery
• Algorithm selection
• Extension
• Implementation
• Results
• Conclusions
• Use decomposed discovery
  – Preserves perfect fitness
• Use only discovery algorithms that guarantee perfect fitness
  – ILP Miner
  – Hybrid ILP Miner
  – Inductive Miner (variant that guarantees perfect fitness)
  – …?

• As a result, all traces of the training logs will be classified as positive
• Use as many decomposed discovery algorithms as are useful
  – An algorithm is considered to be useful if adding it results in additional negatives

• Note that we already have guaranteed that all traces from the training logs will be classified as positive. With adding additional discovery algorithms, we try to squeeze out as many negatives as possible.
• Contest
• Observations
• Model
• Discovery
• Algorithm selection
• Extension
• Implementation
• Results
• Conclusions
• We used the April and May logs to select the best set of useful decomposed discovery algorithms:
  – As few algorithms as needed to achieve the best result
• The best result (showing numbers of traces classified as negative):

<table>
<thead>
<tr>
<th></th>
<th>F</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>March log</td>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
</tr>
<tr>
<td>April log</td>
<td>1</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>9</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>May log</td>
<td>5</td>
<td>10</td>
<td>10</td>
<td>8</td>
<td>10</td>
<td>9</td>
<td>10</td>
<td>10</td>
<td>8</td>
<td>10</td>
<td>10</td>
</tr>
</tbody>
</table>

• As an example, we classified in the May 3 log only 8 traces as negatives. Hence, we know that we have at least 2 false positives.
• The best set of useful decomposed discovery algorithms:
  – Hybrid ILP Miner without decomposition (HIM-0)
  – Inductive Miner with maximal decomposition (IM-100)
Result from the April and May logs, as confirmed by organizers:

- 1 misclassification for the April logs
- 5 misclassifications for the May logs
- These number match the numbers of ‘known’ false positives

<table>
<thead>
<tr>
<th>Training log</th>
<th>F</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>April log</td>
<td>1</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>9</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>May log</td>
<td>5</td>
<td>10</td>
<td>10</td>
<td>8</td>
<td>10</td>
<td>9</td>
<td>10</td>
<td>10</td>
<td>8</td>
<td>10</td>
<td>10</td>
</tr>
</tbody>
</table>

- No false negatives!
• Contest
• Observations
• Model
• Discovery
• Algorithm selection
• Extension
• Implementation
• Results
• Conclusions
• The DrFurby Classifier enriches the log-to-classify, using a “drfurby” extension:
  – Log attributes:
    • name: Name of the log-to-classify
    • positive: Number of traces classified as positive
    • negative: Number of traces classified as negative
    • millis: Number of milliseconds it took to classify the log
  – Trace attributes:
    • classification: “positive” or “negative”
    • him0Costs: Costs of replaying this trace on the net as discovered by IM-100
    • im100Costs: Costs of replaying this trace on the net as discovered by HIM-0
    • totalCosts: Accumulated costs of replaying this trace on all discovered nets
• Observations
• Model
• Discovery
• Algorithm selection
• Extension
• Implementation
• Contest
• Conclusions
The plug-in

Classify with DrFurby Classifier
H.M.W. Verbeek (h.m.w.verbeek@tue.nl)

The inputs
Train-03 XLog

May-03 XLog

The outputs
Test Log with DrFurby Classification XLog

Implementation: DrFurby Classifier plug-in in ProM 6.6 (DivideAndConquer package)
Implementation: result of HIM0 on log 3
Implementation: result of IM100 on log 3
Implementation: result of IM100 on log 3
Implementation: result of IM100 on log 3
Overview

- Contest
- Observations
- Model
- Discovery
- Algorithm selection
- Extension
- Implementation
- Results
- Conclusions
• Results on the June logs:
  – Results for April and May logs added to allow comparison

<table>
<thead>
<tr>
<th>March log</th>
<th>F</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>April log</td>
<td>1</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>9</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>May log</td>
<td>5</td>
<td>10</td>
<td>10</td>
<td>8</td>
<td>10</td>
<td>9</td>
<td>10</td>
<td>10</td>
<td>8</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>June log</td>
<td>7</td>
<td>10</td>
<td>10</td>
<td>7</td>
<td>10</td>
<td>10</td>
<td>9</td>
<td>9</td>
<td>8</td>
<td>10</td>
<td>10</td>
</tr>
</tbody>
</table>

• 7 misclassifications
  – 7 false positives
  – No false negatives
• With 193 correctly classified traces, we won the contest.
• As a result, we received:
  – a free flight to Rio and back (Felix),
  – 4 free nights in a hotel in Rio (Felix),
  – free admission to the conference (Felix), and
  – a trophy, which never materialized (Eric, 😞).
- Contest
- Observations
- Model
- Discovery
- Algorithm selection
- Extension
- Implementation
- Results
- Conclusions
• DrFurby Classifier
  – Takes a training log and a test log
    • Assumes training log to be free of noise
  – Classifies the traces in the test log using the training log
• Some problems with some logs
  – Like 3 and 8
• No false negatives, only false positives
  – No guarantee on the former, however
Questions?
• “DrFurby”
• Runner-ups
• Decomposition
“DrFurby”
- Is pronounced almost identically as “dr. Verbeek” 😊
- Happened years ago, Boudewijn knows the story, he was there.
• Rahi Ghawi: 192 (7 FP, 1 FN)
  – Either (depends on log at hand):
    • Inductive Miner without decomposition, or
    • ILP Miner with maximal decomposition

• Moshe Steiner and Liat Bodaker: 192 (7 FP, 1 FN)
  – Alpha+
  – Model-specific postprocessing
    • Model should classify all traces from March log as positive
• Why decomposition?
  – Results on the April and May logs:

<table>
<thead>
<tr>
<th>March log</th>
<th>F</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>April log</td>
<td>1</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>9</td>
<td>10</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>May log</td>
<td>5</td>
<td>10</td>
<td>10</td>
<td>8</td>
<td>10</td>
<td>9</td>
<td>10</td>
<td>8</td>
<td>10</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>June log</td>
<td>7</td>
<td>10</td>
<td>10</td>
<td>7</td>
<td>10</td>
<td>10</td>
<td>9</td>
<td>9</td>
<td>8</td>
<td>10</td>
<td>10</td>
</tr>
</tbody>
</table>

  – Results on the same logs without using decomposition:

<table>
<thead>
<tr>
<th>March log</th>
<th>F</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>April log</td>
<td>4</td>
<td>10</td>
<td>10</td>
<td>7</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>9</td>
<td>10</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>May log</td>
<td>8</td>
<td>10</td>
<td>10</td>
<td>5</td>
<td>10</td>
<td>9</td>
<td>10</td>
<td>8</td>
<td>10</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>June log</td>
<td>9</td>
<td>10</td>
<td>10</td>
<td>5</td>
<td>10</td>
<td>10</td>
<td>9</td>
<td>9</td>
<td>8</td>
<td>10</td>
<td>10</td>
</tr>
</tbody>
</table>

  – No change for 9 out of 10 logs, but a significant improvement for log 3