Observation Strategies for Event Detection with Incidence on Runtime Verification

Marco Alberti ¹, Pierangelo Dell’Acqua ², and Luís Moniz Pereira ¹

¹CENTRIA, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa Caparica (Portugal)
   m.alberti@fct.unl.pt, lmp@di.fct.unl.pt

²Dept. of Science and Technology, ITN, Linköping University
   Norrköping (Sweden)
   pierangelo.dellacqua@itn.liu.se

RCRA Workshop, Bologna, June 10-11, 2010
Introduction

- Event: observable state in a system
- Event detection is critical in many applications (monitoring, runtime verification, diagnosis, intention recognition, and more)
- Observation may be costly
- Can the number of observations be reduced without losing certainty about event detection?
Outline

1. Event detection in general
2. Sampling problem for natural patterns
3. Algorithms and experimental evaluation
4. Runtime verification of Linear Temporal Logic formulæ
5. Conclusions
Coordinate system

- Set of points where events can happen
- Coordinates can be interpreted as spatial, temporal, or both

**Definition**

A coordinate system $S$ is the cartesian product of $N$ sets, closed under the sum operation.
Patterns

A pattern associates a value taken from a value set $\mathcal{V}$ to (some) points in the coordinate system.

**Definition**

Given a coordinate system $S$, a pattern $\mathcal{P}$ over $S$ is a partial function from $S$ to some value set $\mathcal{V}$.

**Example**

<table>
<thead>
<tr>
<th>$S$</th>
<th>$\mathcal{V}$</th>
<th>$\mathcal{P}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mathbb{N}$</td>
<td>${1}$</td>
<td>${\langle 0, 1 \rangle, \langle 4, 1 \rangle, \langle 5, 1 \rangle}$</td>
</tr>
<tr>
<td>$\mathbb{R}^2$</td>
<td>$\mathbb{R}$</td>
<td>${\langle\langle 0, 0 \rangle, 0 \rangle}$</td>
</tr>
<tr>
<td>$\mathbb{R}^2$</td>
<td>$\mathbb{R}$</td>
<td>${\langle\langle x, y \rangle, x \rangle</td>
</tr>
<tr>
<td>$\mathbb{R}^2$</td>
<td>$\mathbb{R}$</td>
<td>${\langle\langle x, y \rangle, 1 \rangle</td>
</tr>
</tbody>
</table>
Events

- An association of values to points in a coordinate system that follows a given pattern.
- Defined as the translation of its pattern by an offset.

**Definition**

Given a pattern $P$ and $\tau \in S$, the event $E_P(\tau)$ of pattern $P$ and offset $\tau$ is the set $\{\langle \tau + p, P(p) \rangle | p \in \text{dom } P\}$.

- *Unfolding*: all possible events of a given pattern

**Definition**

Given a pattern $P$ in a coordinate system $S$, the unfolding $U_P$ of $P$ is the set $\{E_P(\tau) | \tau \in S\}$.
Observability

Definition
An event e’s **observability set** $\mathcal{O}(e)$ is e’s projection over $S$.

Intuitively, the observability set is the set of points in the coordinate system where the event can be observed.

Definition
A set $\mathcal{C} \subseteq S$ **covers** a pattern $\mathcal{P}$ if and only if for each $\tau \in S$ there exists an element of $\mathcal{O}(\mathcal{E}_\mathcal{P}(\tau))$ that belongs to $\mathcal{C}$.

If a set covers a pattern, the occurrence of any event of that pattern can be decided by observing the set.
Sampling problem

**Definition**

Given a pattern $\mathcal{P}$ over $S$, the **sampling problem** consists of finding a set $\mathcal{C} \subseteq S$ that covers $\mathcal{P}$.

- Trivial solution ($S$), but
- the solution may be required to enjoy some property. For instance:
  - Minimize a cost function (if the cost of observation is a constant, that translates to minimum cardinality)
  - if the points in the coordinate system represent time points, the covering set may be required to intersect observability sets at or near their minimum, to detect events as soon as possible.
Outline

1. Event detection in general
2. Sampling problem for natural patterns
3. Algorithms and experimental evaluation
4. Runtime verification of Linear Temporal Logic formulæ
5. Conclusions
**Natural patterns**

**Definition**

A **natural pattern** is a pattern over the set $\mathbb{N}$ of natural numbers.

Typical interpretation: time points in a discrete time system.

**Example**

Let $\text{dom } \mathcal{P} = \{0, 2, 5\}$. Then the observability set varies with the offset as follows:

<table>
<thead>
<tr>
<th>$\tau=0$</th>
<th>0</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>
<th>11</th>
<th>12</th>
<th>13</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Natural patterns

Definition

A **natural pattern** is a pattern over the set \( \mathbb{N} \) of natural numbers.

Typical interpretation: time points in a discrete time system.

Example

Let \( \text{dom } \mathcal{P} = \{0, 2, 5\} \). Then the observability set varies with the offset as follows:

\[
\begin{array}{cccccccccccccc}
\tau=1 & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 \\
\end{array}
\]
Natural patterns

**Definition**

A **natural pattern** is a pattern over the set \( \mathbb{N} \) of natural numbers.

Typical interpretation: time points in a discrete time system.

**Example**

Let \( \text{dom } \mathcal{P} = \{0, 2, 5\} \). Then the observability set varies with the offset as follows:

\[
\begin{array}{cccccccccccc}
\tau=2 & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 \\
\end{array}
\]
Natural patterns

Definition

A **natural pattern** is a pattern over the set $\mathbb{N}$ of natural numbers.

Typical interpretation: time points in a discrete time system.

Example

Let $\text{dom } \mathcal{P} = \{0, 2, 5\}$. Then the observability set varies with the offset as follows:

<table>
<thead>
<tr>
<th>$\tau=3$</th>
<th>0</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>
<th>11</th>
<th>12</th>
<th>13</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Natural patterns

Definition

A **natural pattern** is a pattern over the set \( \mathbb{N} \) of natural numbers.

Typical interpretation: time points in a discrete time system.

Example

Let \( \text{dom } \mathcal{P} = \{0, 2, 5\} \). Then the observability set varies with the offset as follows:

<table>
<thead>
<tr>
<th>( \tau = 4 )</th>
<th>0</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>
<th>11</th>
<th>12</th>
<th>13</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Natural patterns

Definition

A natural pattern is a pattern over the set $\mathbb{N}$ of natural numbers.

Typical interpretation: time points in a discrete time system.

Example

Let $\text{dom } \mathcal{P} = \{0, 2, 5\}$. Then the observability set varies with the offset as follows:

<table>
<thead>
<tr>
<th>$\tau=5$</th>
<th>0</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>
<th>11</th>
<th>12</th>
<th>13</th>
</tr>
</thead>
</table>
Normalized patterns

**Definition**

Given a natural pattern $\mathcal{P}$, let $m \triangleq \min \text{dom } \mathcal{P}$.

Then the **normalization** of $\mathcal{P}$ is the pattern $\overline{\mathcal{P}} \triangleq \{\langle p, \mathcal{P}(p + m) \rangle \mid p + m \in \text{dom } \mathcal{P}\}$.

**Example**

Let $\text{dom } \mathcal{P} \triangleq \{8, 10, 13\}$. Then $\text{dom } \overline{\mathcal{P}} = \{0, 2, 5\}$.

**Theorem**

*Let $\mathcal{P}$ be a natural pattern. Then an event of pattern $\mathcal{P}$ is also an event of pattern $\overline{\mathcal{P}}$.***
Circular repetition

**Definition**

Let $\mathcal{I}$ be a finite set of natural numbers, and let $M \triangleq \max \mathcal{I}$. For an integer $k$, let $\sigma_k \triangleq \{(p + k) \mod (M + 1) \mid p \in \mathcal{I}\}$.

$\mathcal{I}$’s **circular repetition** is the collection of sets $\mathcal{U} \triangleq \{\sigma_k \mid k \in [0 .. M]\}$.

**Example (cont’d)**

$\text{dom } \overline{P}$’s circular repetition is $\mathcal{U} \triangleq \{\sigma_0, \ldots, \sigma_5\}$, where

- $\sigma_0 \triangleq \{0, 2, 5\}$,
- $\sigma_1 \triangleq \{0, 1, 3\}$,
- $\sigma_2 \triangleq \{1, 2, 4\}$,
- $\sigma_3 \triangleq \{2, 3, 5\}$,
- $\sigma_4 \triangleq \{0, 3, 4\}$, and
- $\sigma_5 \triangleq \{1, 4, 5\}$.
Covering shape

Definition

Let $\mathcal{P}$ be a natural pattern. Then a set with non-empty intersection with each element of the circular repetition of $\text{dom} \ \mathcal{P}$ is a covering shape of $\mathcal{P}$’s.

Example (cont’d)

A (minimal) covering shape is $\mathcal{S}_C \triangleq \{0, 1, 3\}$.
From covering shape to covering set

**Theorem**

Let $P$ be a natural pattern and let $M \triangleq \text{max dom } P$. Let $S_C$ be a covering shape of $P$’s. Then the set $C \triangleq \{q + k(M + 1) \mid q \in S_C \land k \in \mathbb{N}\}$ covers $P$.

**Example (cont’d)**

The corresponding covering set of $P$’s is $C \triangleq \{q + 6k \mid q \in \{0, 1, 3\} \land k \in \mathbb{N}\} = \{0, 1, 3, 6, 7, 9, 12, \ldots\}$, which also covers $P$. 

\[0 \quad 1 \quad 2 \quad 3 \quad 4 \quad 5 \quad 6 \quad 7 \quad 8 \quad 9 \quad 10 \quad 11 \quad 12 \quad 13\]
Summary so far

- The sampling problem for a natural pattern $\mathcal{P}$ can be reduced to finding a set that intersects a collection of $M + 1$ subsets of $[0 .. M]$, where $M \triangleq \max \text{dom } \mathcal{P}$.
- The periodic repetition of such set, with period $(M + 1)$, covers the pattern.
- A trivial, and uninteresting, solution is $[0 .. M]$ itself (which corresponds to observing the system at all time points).
- However, finding such a set with minimum cardinality is a well known NP-complete problem: the minimum hitting set problem (given a collection of sets, find a minimal set with non-empty intersection with all of them).
Sampling ratio

How many observations does reduction save us?

**Definition**

Given a normalized pattern $\mathcal{P}$ and a covering shape $S_C$ of $\mathcal{P}$’s, the corresponding **sampling ratio** is

$$R_{S_C, \mathcal{P}} \triangleq \frac{|S_C|}{1 + \text{max dom } \mathcal{P}}.$$ 

**Example (cont’d)**

The sampling ratio is $R_{S_C, \mathcal{P}} = 3/6 = 1/2$.

**Example**

| $|\text{dom } \mathcal{P}|$ | 4  | 8  | 12 | 16 | 20 |
|--------------------------|----|----|----|----|----|
| $R_{S_C, \mathcal{P}}$    | 0.32 | 0.19 | 0.14 | 0.1 | 0.1 |

Average sampling ratios computed on random patterns (max dom $\mathcal{P} = 20$).
If the domain is a natural interval

- Let a pattern $\mathcal{P}$’s domain be $[0 .. M]$.
- $\text{dom} \ \mathcal{P}$’s circular repetition is $\mathcal{U} \triangleq \{\sigma_0, \ldots, \sigma_M\}$, where for $i \in [0 .. M]$ $\sigma_i = [0 .. M]$.
- A minimum hitting set for $\mathcal{U}$ is $\{i\}$, for any $i \in [0 .. M]$.
- For each $i$, a covering set is $\{i + k(M + 1) \mid k \in \mathbb{N}\}$.
- The corresponding sampling ratio is $\frac{1}{M+1}$.
- Intuitive: if an event lasts $M + 1$ time points, it is sufficient to observe every $(M + 1)$-th point to detect its occurrence.
Outline

1. Event detection in general
2. Sampling problem for natural patterns
3. Algorithms and experimental evaluation
4. Runtime verification of Linear Temporal Logic formulæ
5. Conclusions
Complete algorithm

- Based on a total order of all the elements of $2^{[0..\text{max dom} \mathcal{P}]}$.
- One is returned as the minimum hitting set.
- As a first attempt, the full integer interval from 0 to $\text{max dom} \mathcal{P}$, which obviously hits the pattern domain’s circular repetition.
- At each iteration, it tests for hitting the next set with lesser cardinality than the current best, according to the aforementioned order.
Reduction to set cover and approximated solution

Each minimum hitting set problem can be transformed into a set cover problem: given a collection \( S \subseteq 2^{[1 \ldots n]} \), find a subset of \( S \) of minimal cardinality whose union is \([1 \ldots n]\).

- Well studied: performance (ratio between cardinality of computed cover and minimal cover) bounds for approximation algorithms.
- Greedy algorithm:
  - At each step, choose the biggest set, and delete its elements from the other sets, until the universe is covered.
  - Performance ratio is close to theoretical lower bound for polynomial algorithms (unless \( P = NP \)).
## Experimental comparison

| $|\text{dom } \mathcal{P}|$ | Ratio (average) | Ratio (max) | Time (complete) | Time (greedy) |
|----------------|----------------|--------------|----------------|--------------|
| 2              | 1.0            | 1.0          | 24777.18       | 2.491        |
| 4              | 1.033          | 1.333        | 2023.496       | 1.915        |
| 6              | 1.02           | 1.2          | 296.656        | 1.83         |
| 8              | 1.0            | 1.0          | 48.383         | 1.73         |
| 10             | 1.033          | 1.333        | 21.473         | 1.787        |
| 12             | 1.0            | 1.0          | 7.937          | 1.716        |
| 14             | 1.0            | 1.0          | 6.318          | 1.675        |
| 16             | 1.0            | 1.0          | 1.556          | 1.614        |
| 18             | 1.0            | 1.0          | 1.363          | 1.843        |
| 20             | 1.0            | 1.0          | 1.522          | 1.668        |

Results on random patterns (50 iterations)

$\max \text{dom } \mathcal{P} = 20$, theoretical lower bound is 1.62.
Outline

1. Event detection in general
2. Sampling problem for natural patterns
3. Algorithms and experimental evaluation
4. Runtime verification of Linear Temporal Logic formulæ
5. Conclusions
Sampling of paths in Kripke structures

Definition

Given a Kripke structure $\mathcal{K}$, a path $\pi \triangleq s_0s_1 \ldots s_n \ldots$ on $\mathcal{K}$, and a finite set $\mathcal{I}$ of natural numbers whose maximum is $M$, $\pi$’s **sampling** of shape $\mathcal{I}$, $\pi_\mathcal{I}$, is the sequence of states $s_i$ such that $s_i$ is a state in $\pi$ and $i = q + k(M + 1)$, for some $k \in \mathbb{N}$ and $q \in \mathcal{I}$.

Example

The sampling of a path $\pi \triangleq s_0s_1 \ldots s_{16}$ of shape $\{0, 1, 3, 6\}$ is $s_0s_1s_3s_6s_7s_8s_{10}s_{13}s_{14}s_{15}$.
LTL events

Definition

Given a natural pattern $\mathcal{P}$, a formula $\varphi$ is an **LTL-event** of pattern $\mathcal{P}$ in a Kripke structure $\mathcal{K}$ if and only if for any path $\pi \triangleq s_0s_1\ldots s_n$ in $\mathcal{K}$,

$$(\exists k \mid s_k \models \varphi) \Rightarrow \exists \tau \ (\exists l \in \text{dom } \mathcal{P} \mid \tau + l = k \land \forall j \in \text{dom } \mathcal{P} \mid s_{\tau+j} \models \varphi)$$

Example

Let $\text{dom } \mathcal{P} = \{0, 2, 5\}$ and let $\varphi$ be an event of pattern $\mathcal{P}$. Then a possible distribution of truth values for $\varphi$ over the states can be as follows (blue circle means true):

<table>
<thead>
<tr>
<th>$s_0$</th>
<th>$s_1$</th>
<th>$s_2$</th>
<th>$s_3$</th>
<th>$s_4$</th>
<th>$s_5$</th>
<th>$s_6$</th>
<th>$s_7$</th>
<th>$s_8$</th>
<th>$s_9$</th>
<th>$s_{10}$</th>
<th>$s_{11}$</th>
<th>$s_{12}$</th>
<th>$s$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\tau = 3$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Detection of LTL events

**Theorem**

Let $\varphi$ be a LTL-event of pattern $\mathcal{P}$ on a Kripke structure $\mathcal{K}$. Let $S_C$ be a covering shape of $\mathcal{P}$’s. Then, for each path $\pi \triangleq s_0 s_1 \ldots s_n \ldots$ on $\mathcal{K}$, $\pi \models \Diamond \varphi$ if and only if $\pi_{S_C} \models \Diamond \varphi$.

Reduction of observation points in runtime verification of LTL formulae, in case of event detection.
Outline

1. Event detection in general
2. Sampling problem for natural patterns
3. Algorithms and experimental evaluation
4. Runtime verification of Linear Temporal Logic formulæ
5. Conclusions
Summary of results

- General formulation of sampling problem for event detection
- For natural patterns, reduction to (NP-complete) minimum hitting set (approximation algorithms available)
- Application to runtime verification of LTL formulæ
Generalizations

- Multi-dimensional natural coordinate systems.
- Different patterns in the same system:
  - Disjoint value sets (events immediately recognized as belonging to a pattern);
  - Intersecting value sets (more observations necessary).
- If event detection is not critical, further reduce number of observations (decision theory).
Beyond simple event detection

Representation of relationships among events:

- Event algebras
- Integrity constraints to define permissible combinations
- Tell-tale signs to direct attention
- Probabilities (Bayesian networks)
- More than one possible solution:
  - Multi-threaded computations
  - Preferences