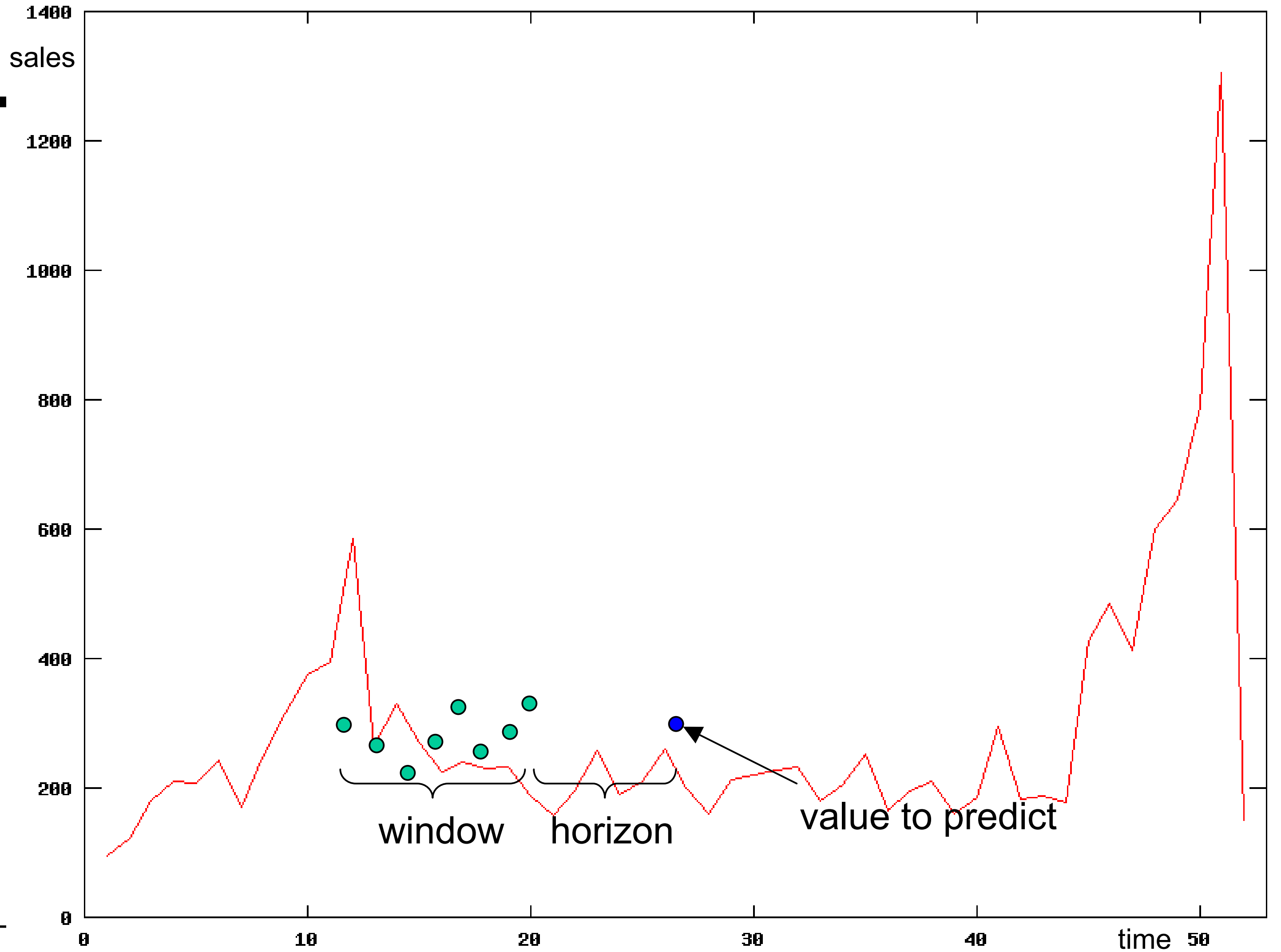
Summarizing Time Series and the Detection of Event Sequences

Katharina Morik
Univ. Dortmund

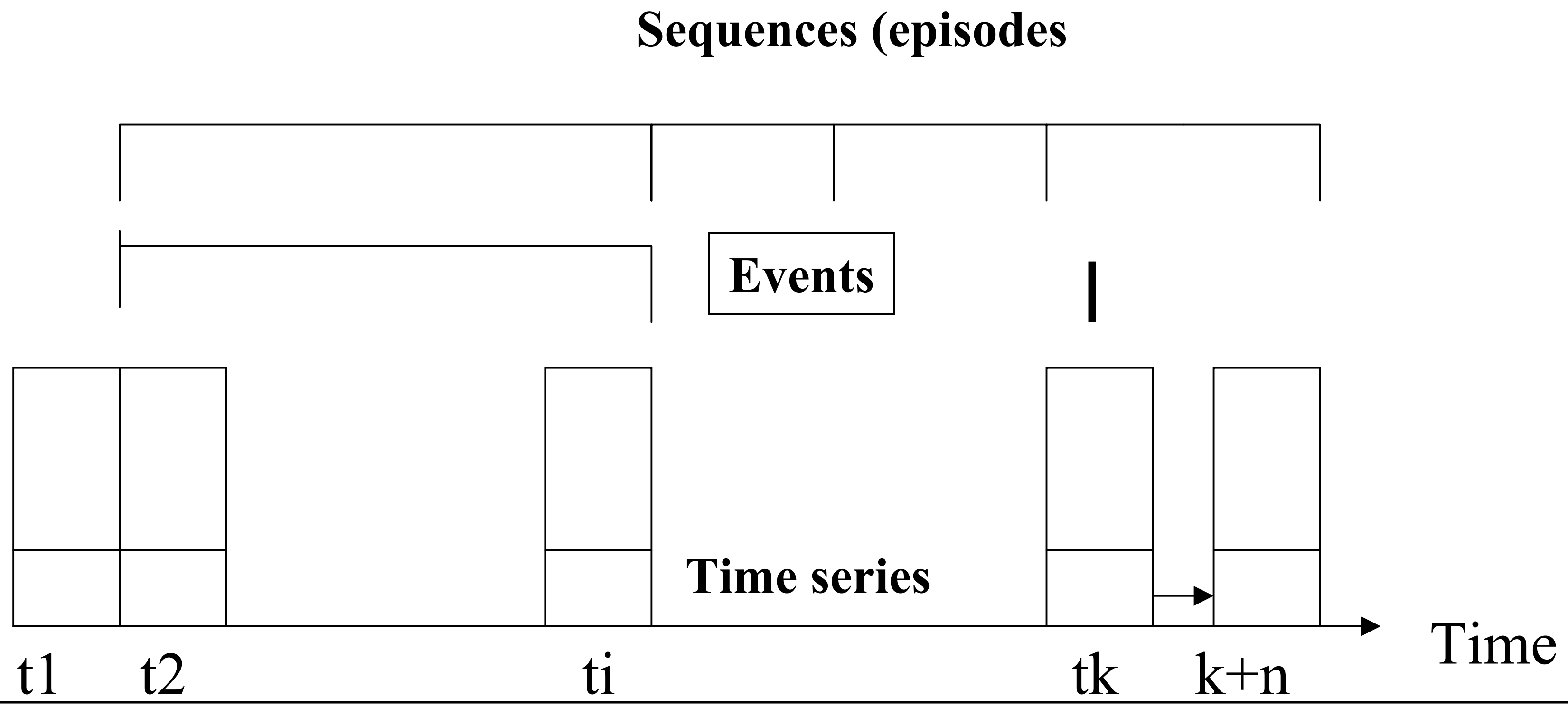
Learning Tasks
Time Series Analysis

- Observations are produced by 1 stable process.
  Given k observations,
  – predict the k+n observation -- forecasting
  – detect a trend, a season, a cyclic behavior
  – find outliers

Forecasting Sales



Time Phenomena



Substrings

If events must occur contiguously, it is a string.

String: abchhttaabcttababcbbhtt

Substrings: abchhttaabcttababcbbhtt

...

abc aab aba bch bct bab bcb bbh bht chh cbb ctt hht htt tta taa tab

...

ab aa ab bc ba bb bh ch ct hh ht tt ta

a, b, c, h, t

Boyer-Moore algorithm for deciding whether s ist substring of s*

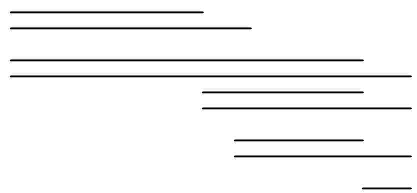
Subsequences

In a subsequence, events might be interspersed with other events not in the subsequence.

String: abchhttaabcttababcbbhtt

Subsequences:

abchhttaabcttababcbbtt



Frequent subsequences: All t come as substring `tt´.

a*a is a frequent subsequence.

a*c*a is a frequent subsequence...

Learning Tasks
From Series to Sequence

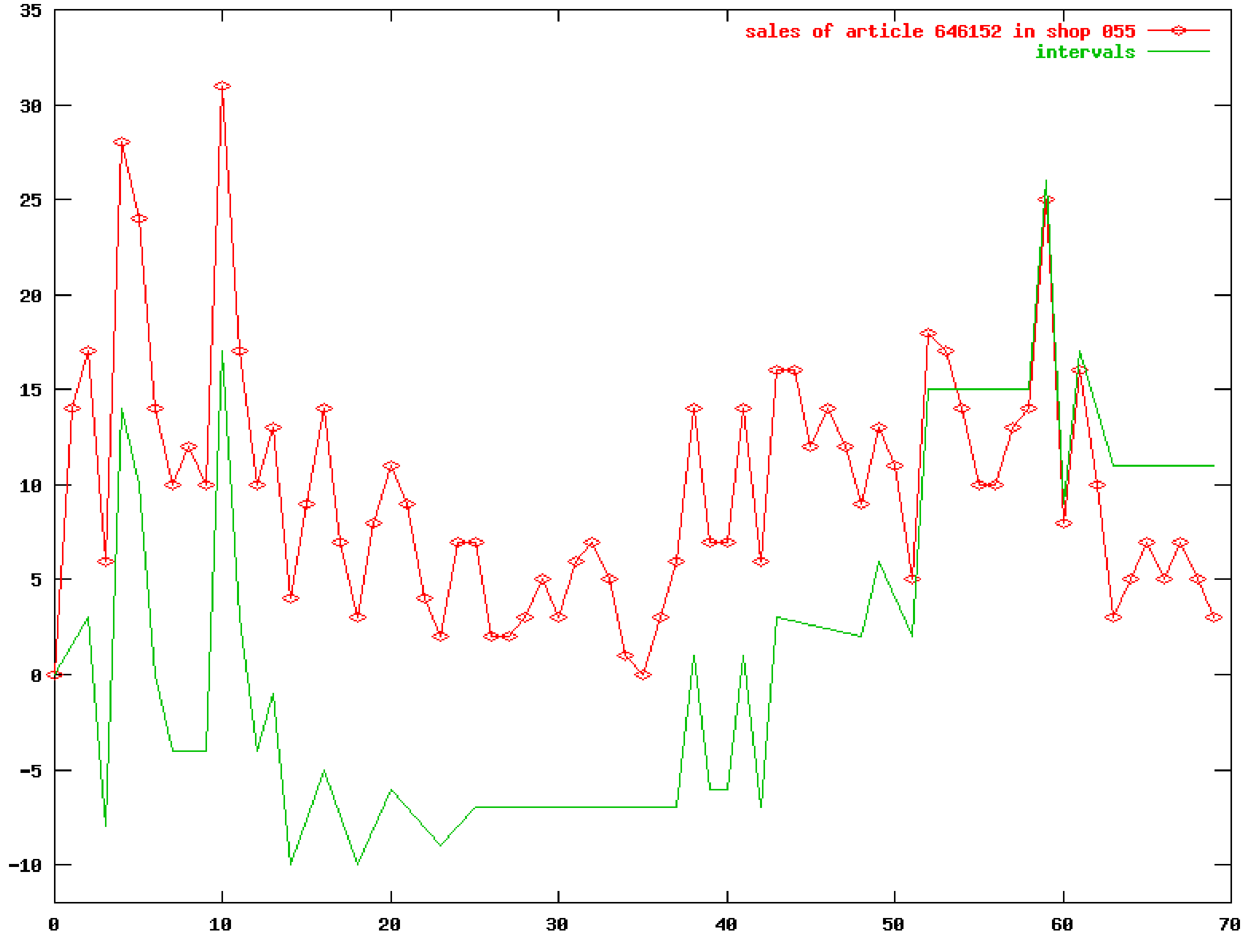
- Given some time series
  - detect events (states, intervals) – abstraction

An event is a triple (state, begin,finish).

The state might be a label or a (mean) value.

Typical labels are: increase, decrease, stable...

Summarizing Item 646152 in Shop 55



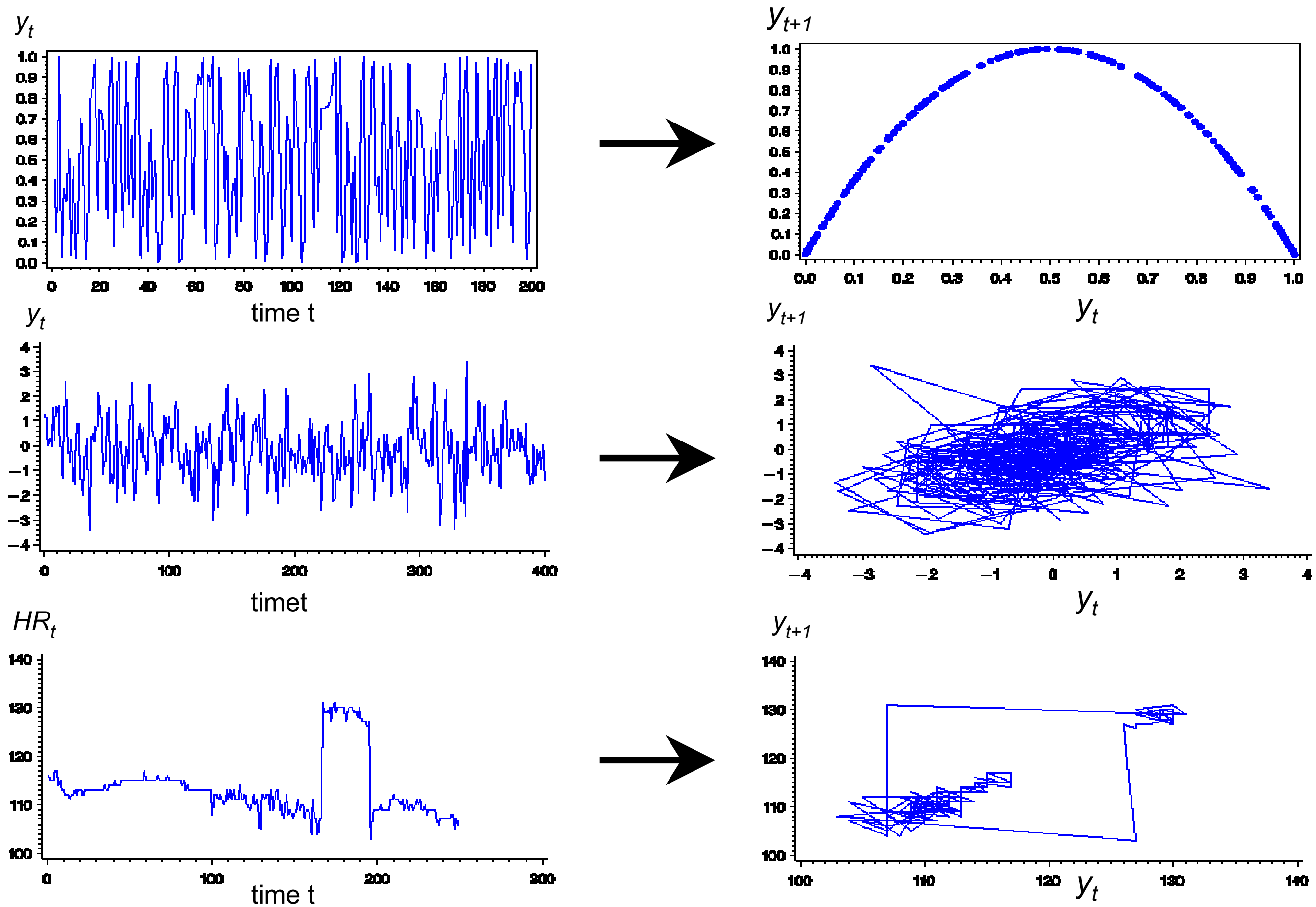
Corresponding Facts

small time intervals

increasing(646152,1,2,3).
decreasing(646152,2,3,-11).
increasingPeak(646152,3,4,22).
...
stable(646152, 25,37,0).
increasing(646152, 37, 38, 8).
decreasing(646152, 38, 39, -7).
stable(646152, 39,40, 0).
increasing(646152, 40, 41,7).
decreasing(646152, 41, 42,-8).
increasing(646152, 42, 43,10).
stable(646152, 43, 48,-1).

Phase State Analysis

Time series $y_1, \dots y_N$ **Phase state** $\vec{y}_t = (y_t, y_{t+1})$



U.Gather, M. Bauer

Unsupervised Methods

- Statistical discovery of level changes – all contiguous observations within one level form one event (Markus Bauer).
- Statistical discovery of trends – all contiguous observations with more or less the same gradient form one event (Morik, Wessel).
- Clusters of subsequences form events (Das et al.).

Approaches Presented at the WS

- Fu-Lai Chung defines defines critical points and forms typical shapes which can be used for determining the similarity of different time series.
- Edwin Pednault considers quarterly intervals a state and models the distribution of events within states.
- Ralf Klinkenberg investigates sequences of document classifications and determines where the process changes using the epsilon-alpha estimates of the SVM.

Learning Tasks
Discovery of Frequent Patterns

- Given a large set of processes with their observations,
  - find (dis-)similarities among the time series
  - find frequent subsequences
  - find frequent substrings
- Given a set of processes with their events,
  - find temporal or causal relations between events

Relations Between Events

- Events may overlap, meet, occur at the same time, follow one another, one may start or finish the other – Allen‘s 13 relationships.

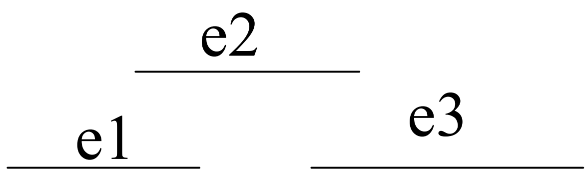
e1(Begin1, End1) e2(Begin2, End2)

e3(Begin3, End3)

Begin1< Begin2   Begin2 < End1

End1 < End2   Begin3 < End2

End1 < Begin3



Rule Learning

- Rules about sales trends:

increasing (Item, T_b, T_e), decreasing(Item, T_e, T_{e2})
$\rightarrow$stable(Item, T_{e2}, T_{e3})

Using the Rule Discovery Tool (Kietz,Wrobel) of inductive logic programming, which can also learn relations beween events of different agents (processes).

Use of Temporal Relations Presented at the WS

- Ursula Sondhauß uses 4 economical phases as states and models the pairwise transition probability where illegal transitions are excluded.
- Malek Mouhoub combines several temporal relations between events of several subjects into valid logic models.
- Frank Höppner forms patterns out of temporal relationships and detects the most frequent ones.