Tutorial at Measuring Behavior 2010

Analyzing Behavior and Interactions with

THEME™

Detection and Analysis of Hidden Temporal Patterns

Magnus S. Magnusson
Research Professor, director
Human Behavior Laboratory
University of Iceland
www.hbl.hi.is

About Notation

- The prefix “T-”, as in T-pattern, below only indicates that the prefixed term is used in a particular way specified through a formal definition within the evolving T-system or T-language for the description and analysis of spatial and/or temporal structure.

A Long Search for Patterns

First in Human Interactions, then in neurons and DNA

- This research and development work has been carried out at the Universities of Copenhagen, Paris (V, VIII & XIII) and Iceland and at the Museum of Natural History in Paris.
- There has been extensive collaboration with research teams in a number of universities in the USA and Europe, especially the University of Chicago and the MASI university network that was created around the T-pattern model and the use of the THEME software .

- See www.hbl.hi.is and http://www.hbl.hi.is/masi.htm

- From the beginning of the present study in the late 70's the focus has been on formulating and discovering some general structural aspects of behavioral dynamics as a necessary prerequisite for the subsequent development of adequate generative models – still out of reach.

Why Search For Repeated Patterns?

Repeated Patterns are Fundamental -- Often Hidden --Biological and Behavioral Phenomena:

"Another key feature of biology is the existence of many identical examples of complex structures."

(Francis Crick, 1989, p. 138; co-discoverer of DNA)

"**Behavior consists of patterns in time**. Investigations of behavior deal with sequences that, in contrast to bodily characteristics, are **not always visible**."

The opening words of Eibl-Eibesfeldt's book:"Ethology: The Biology of Behavior", 1970, p. 1; {Emphasis added.}

What Cannot Be Detected Cannot Be:

– Counted

– Interpreted

– Compared or

– Modeled

For the Discovery of Experimental Effects

- After a particular treatment the absolute and relative frequencies of components may not change at all, while the way they combine into patterns may change dramatically.

- Examples:

- 1) A solution with DNA before and after heating.

- 2) Text before and after some reordering of its letters or words.

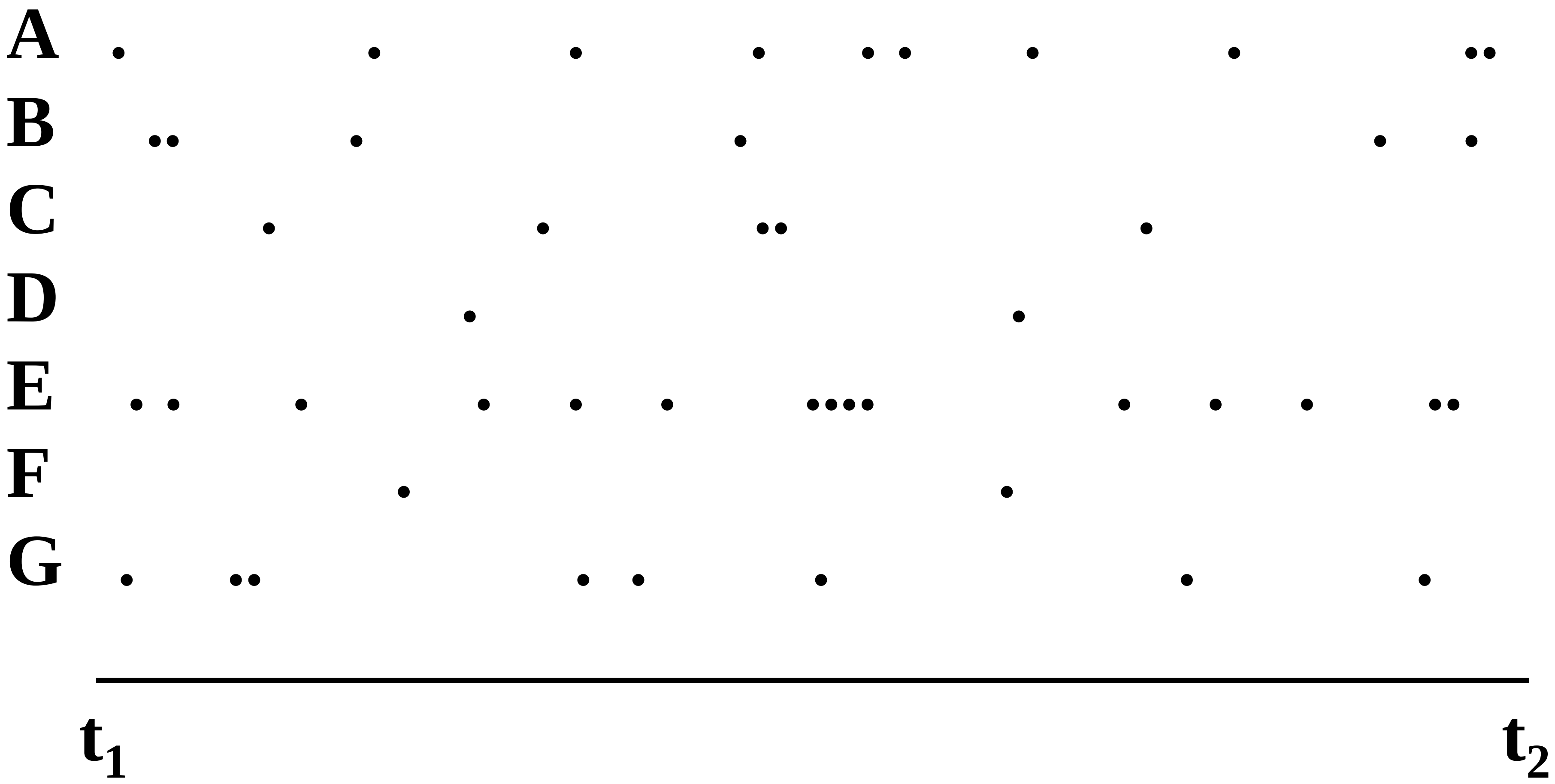
- Detection of these effects requires discovery of patterns before and after treatment. Pattern changes may be the only, but sometimes very strong effects of an independent variable; effects often completely overlooked when pattern detection is omitted.

The Data Type: **T-Data**

- **T-Data** stands for a set of (time) point series where each series (S) stands for the occurrence points of an event-type within k intervals or periods $[\ t_1, t_2\ ]_{1..k}$
- An event-type name, E, (for example, joe,begins,running) only serves as a label for its occurrence series.
- Thus T-Data with n series and k periods can be noted as: $E_i S_i[\ t_1, t_2\ ]_j$; $i=1..n$, $j=1..k$

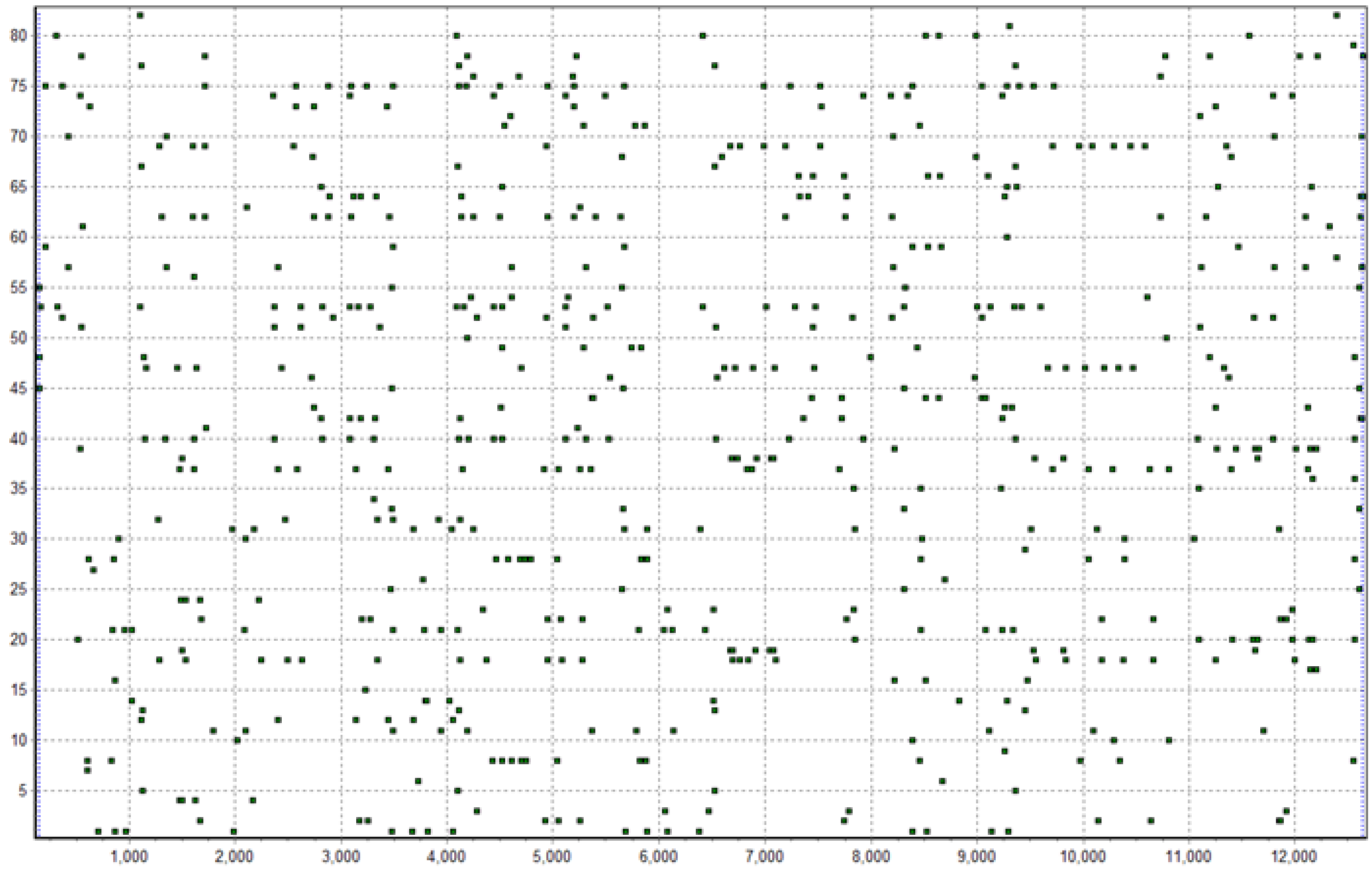
T-Data: Multivariate Point Series

For example, tens, hundreds or thousands of series



Human Interaction T-Data

82 Event-Types, approx. 40 per child, 13:30 min



Time (one frame)1/15 s)

Data Entry and Transformations

- The PatternVision **D**ata **E**xchange **P**rogram: **DEP**
  - Accepts data from most versions of **The Observer** and returns data ready for analysis with **THEME**.
- Analog Data → T-Data
  - Ready soon: a simple separate program for basic transformation of analog data into T-Data

What Kinds of Patterns?

- Do all these phenomena share a pattern type?

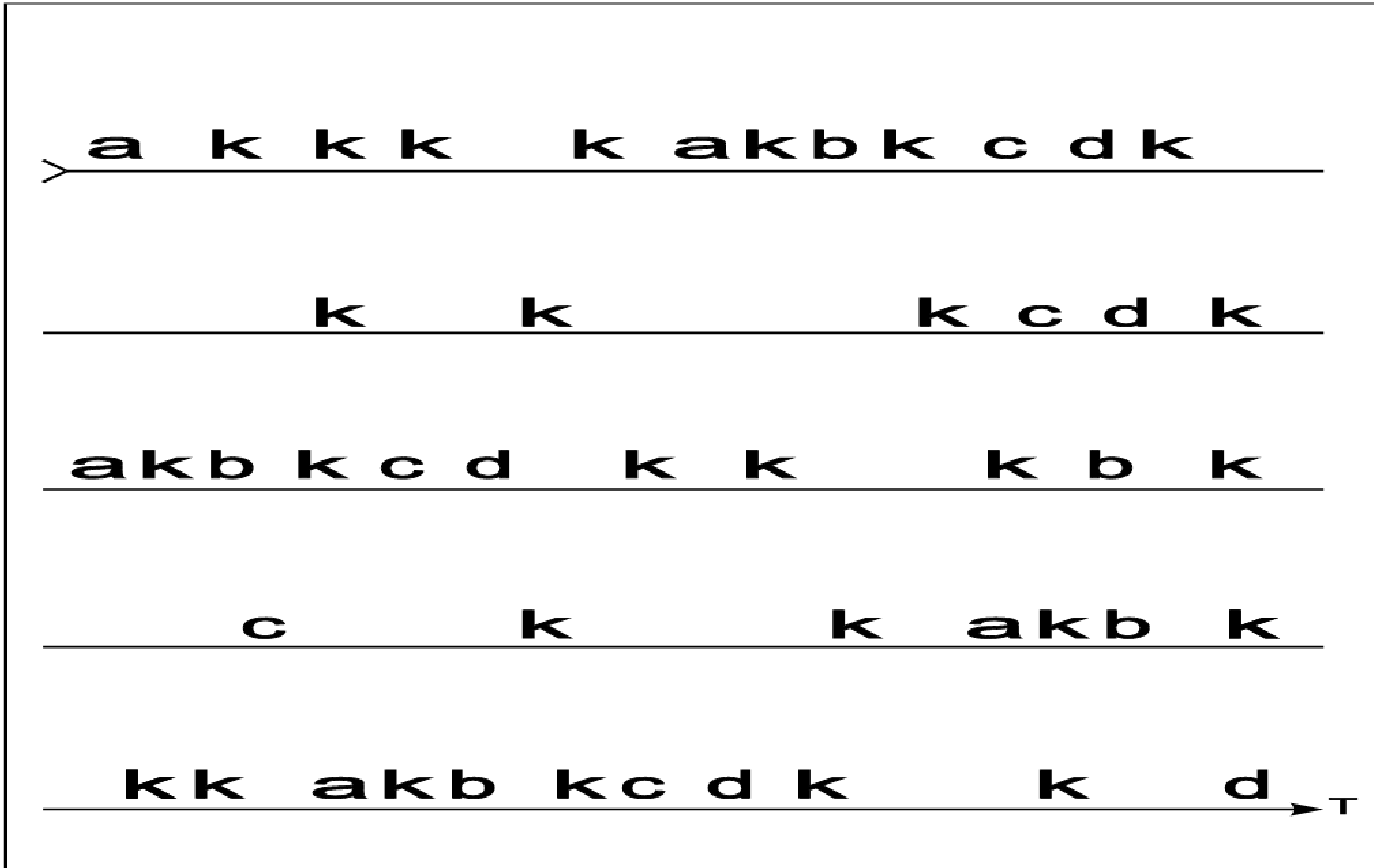
- **DNA, RNA**: Motives, exons, introns, genes…
- **Language**: phonemes, syllables, words, standard phrases, standard texts (poems, scriptures, law books, etc.)
- Nonverbal Behavior and Interactions
- Neuronal Interactions

T-Patterns

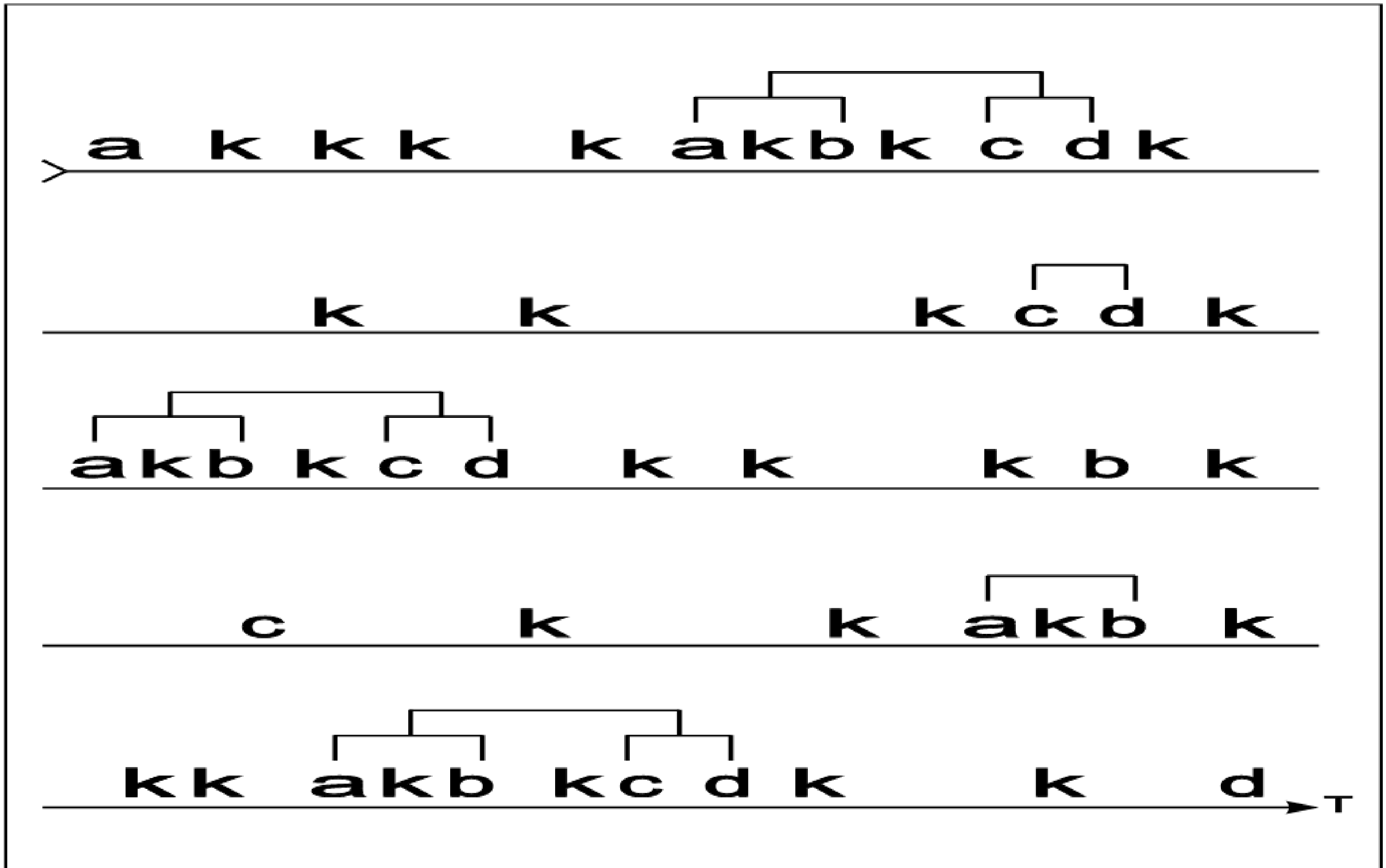
- The T-pattern is the initial and basic pattern type in the T-system.
- It can be described or defined verbally in a generally useful manner, but also formally including algorithmic definitions.
- Thus a T-pattern stands for a set of components that occur repeatedly in the same order such that the time distance between each consecutive pair remains significantly similar, while each component may itself be a T-pattern.

A 1-D Record with Patterns

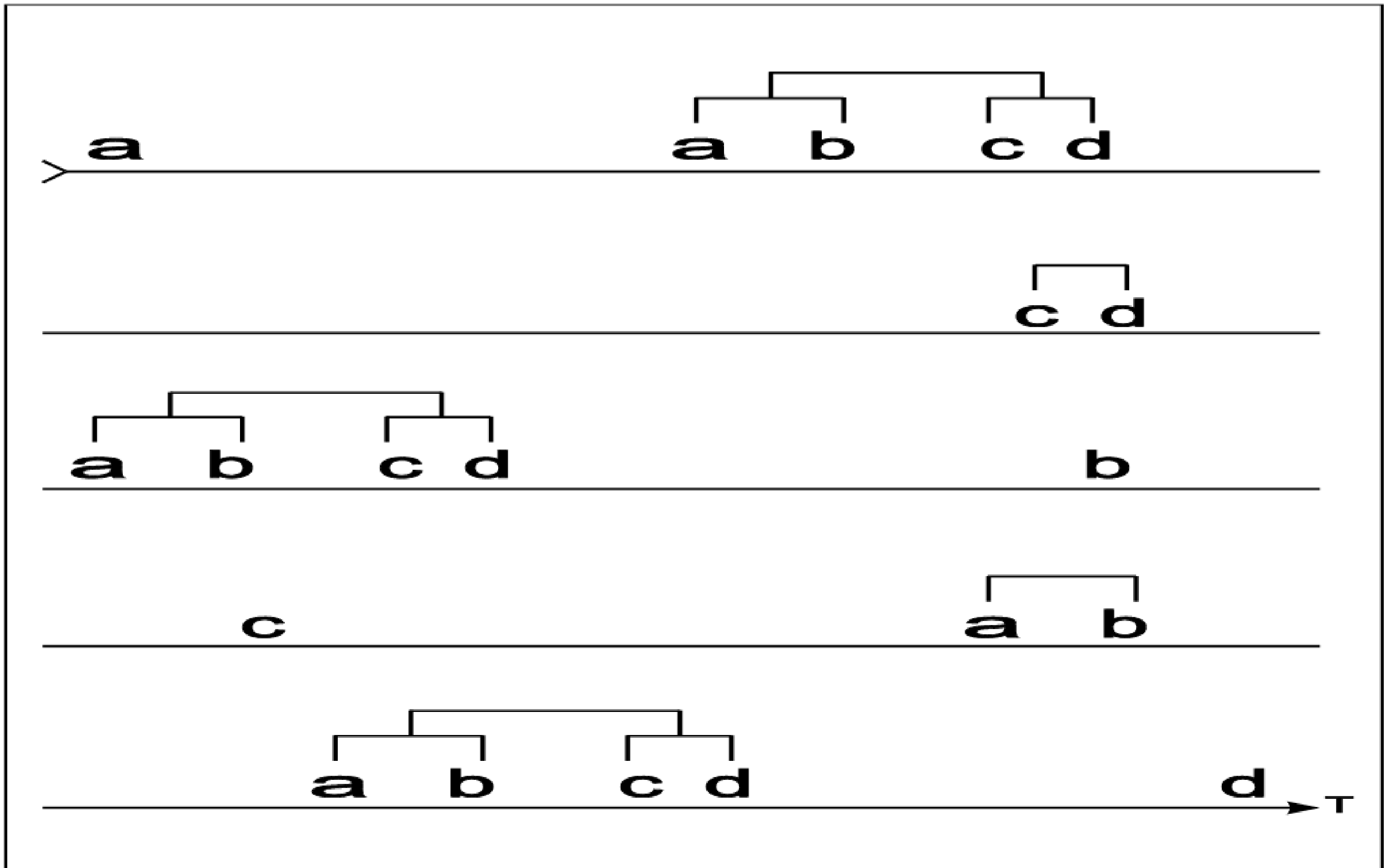
Minimal "Noise" But Pattern Hard to See



...binary trees show the patterns



..noice (K) removed



T-Pattern Definition

With the Initial T-Data Event-Types as the Simplest T-patterns

Within [1, T] a T-pattern is an ordered set of T-patterns (X):

$$X_1 \approx dt_1\ X_2 \approx dt_2\ ..\ X_i \approx dt_i\ X_{i+1}\ ..\ X_{m-1} \approx dt_{m-1}\ X_m$$

that recurs such that each of the time distances $\approx dt_i$ varies significantly less than expected assuming a zero hypothesis (fiction) of constant probability per unit time for each X_i given by

$$X_i = N_{Xi} / T.$$

Thus after X_i occurring at t, X_{i+1} occurs (at least once) within more of the intervals [t+d$_1$, t+d$_2$]$_l$ than expected by chance:

$$X_1\ [\ d_1, d_2\ ]_1\ X_2\ ..\ X_i\ [\ d_1, d_2\ ]_i\ X_{i+1}\ ..\ X_{m-1}\ [\ d_1, d_2\ ]_{m-1}\ X_m$$

Critical Intervals and Binary Trees

Any sequence can be split into two shorter ones

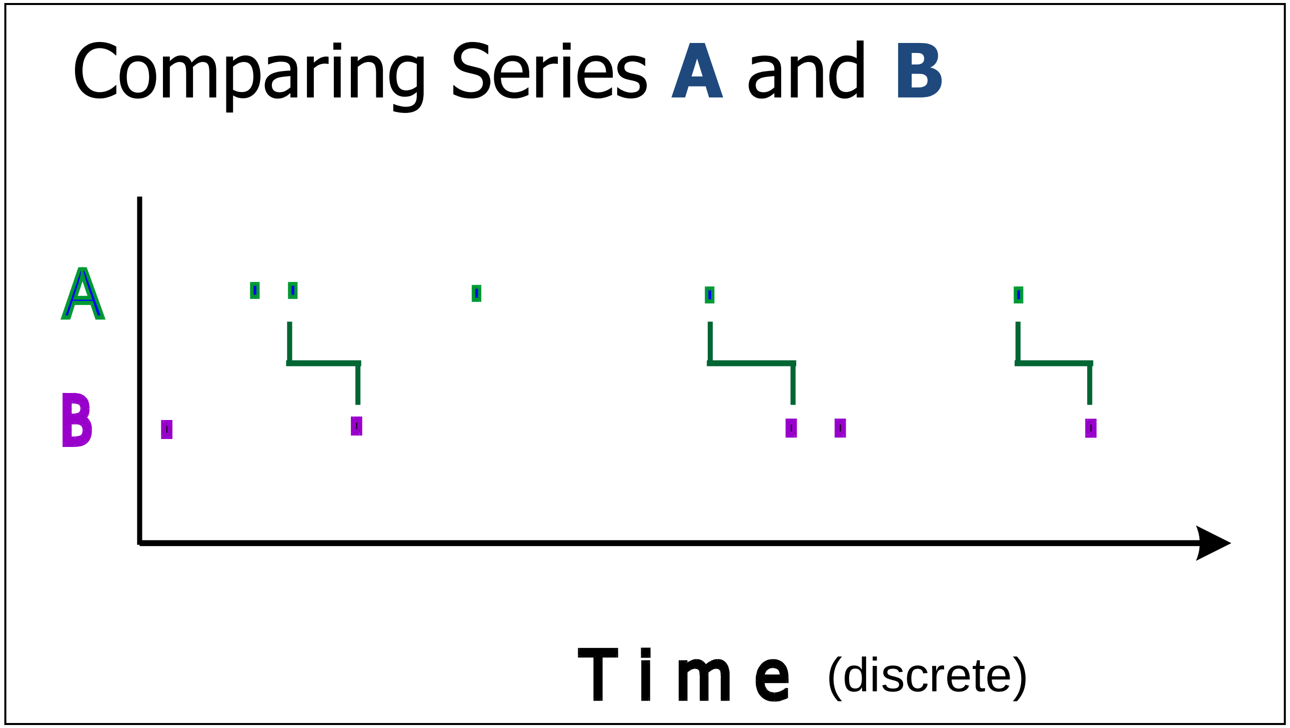
- Any T-pattern $\mathbf{Q = X_1\,X_2..X_m}$ can be split into a **pair** of shorter ones related by a **critical interval**:

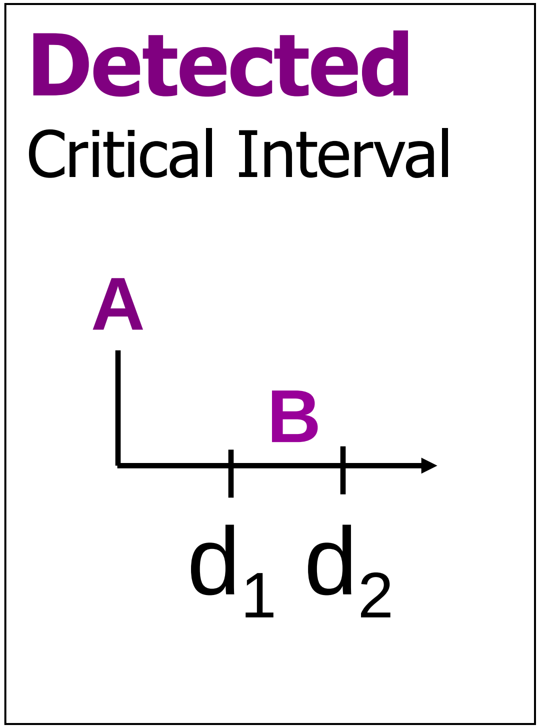
$$\mathbf{Q_{Left}\;[\;d_1\,,\;d_2\;]\;Q_{Right}}$$

- **Recursively**, $\mathbf{Q_{Left}}$ and $\mathbf{Q_{Right}}$ can thus each be split until the pattern $\mathbf{X_1..X_m}$ is expressed as the terminals of a **binary-tree. --** Detection works in the opposite direction.

Towards a Detection Algorithm

T-pattern evidence: Critical Intervals [d_1 , d_2]



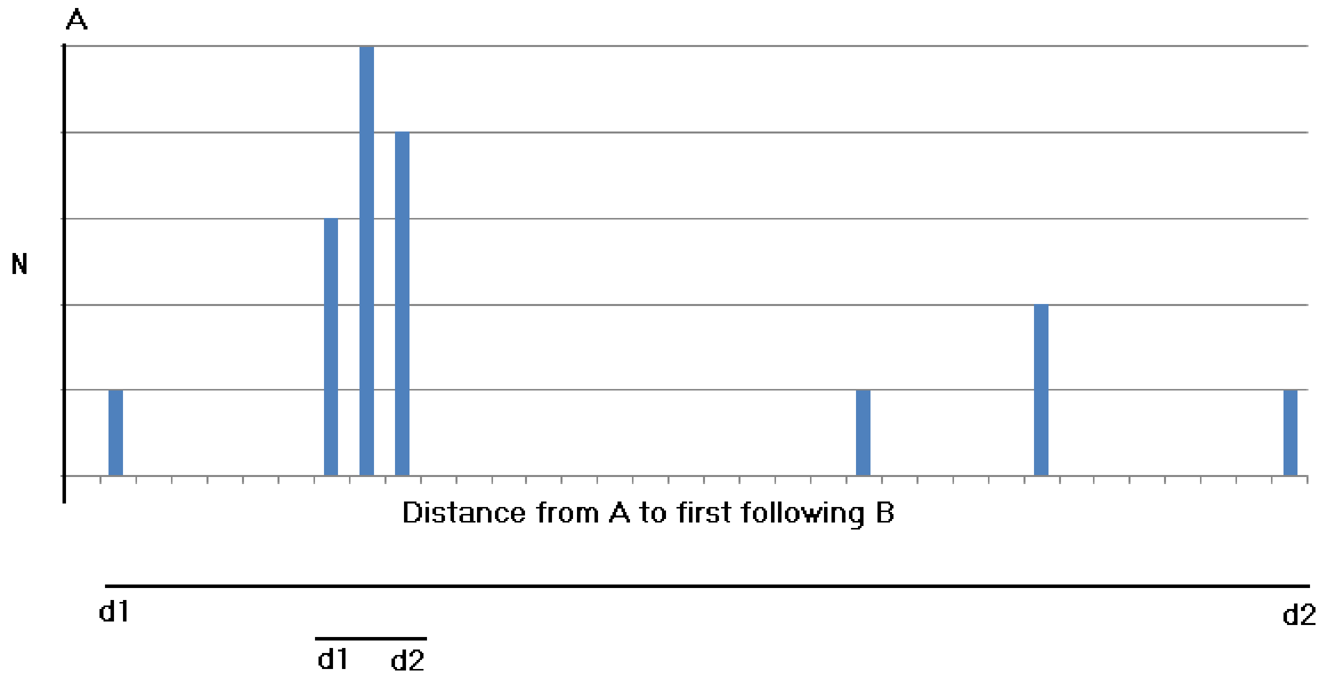


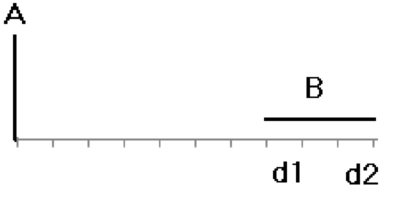
Repeatedly, an A may be followed by a B within approximately the same distance

Search for a critical interval [t+d1, t+d2] after A occurring a t, such that significantly more of these than expected by chance contain at least one occurrence of B?

Searching for [d_1, d_2] in Distributions

Using a Simple Binomial Test





For details see: http://dl.dropbox.com/u/1983864/Magnusson_2000_BRMIC.pdf

Detection as Evolution

Patterns Grow and Compete; The Most Complete Win

- The bottom-up algorithm detects patterns gradually from simpler to more complex as critical pairs of pairs, i.e., as binary trees of critical intervals

- Starting with the coded event types, it detects critical interval relations between the occurrence series of event types and/or detected patterns

- And when found these are connected to form longer patterns (binary trees) which are added to the data (multi ordinal)

- Many binary-trees may correspond fully or partly to the same pattern so all detected patterns are automatically compared and only the most complete (longest) patterns survive.

Completeness Competition

Deletes Partial and Later Detected Equivalent Trees

Same begin and end points and thus the same continuations:

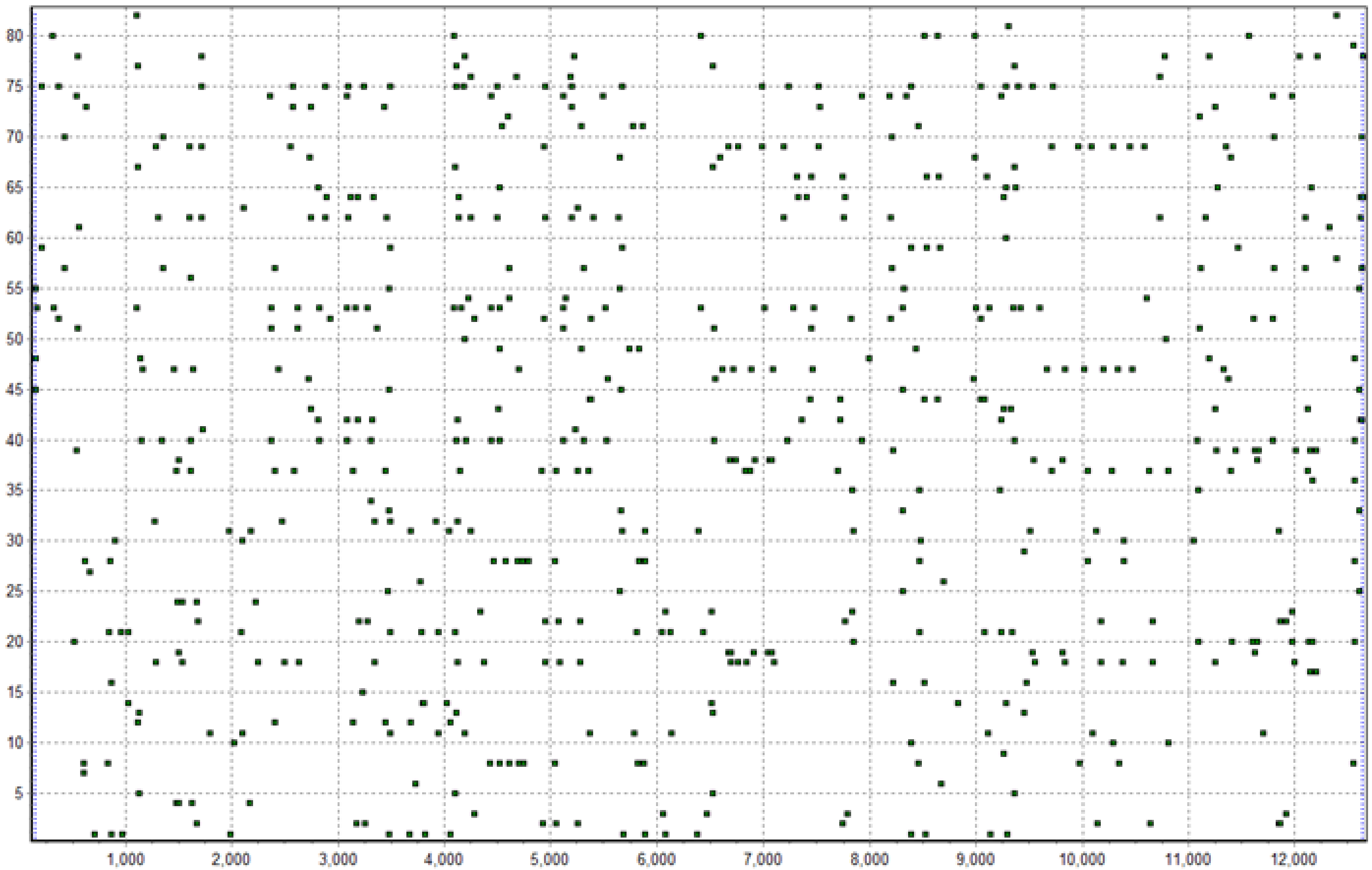
- **Equivalent; contain same event occurrences:**
  - **((A B) ((C D) (E F)))**
  - **((A ((B C) D)) (E F))**
- **Partial; have some of the same event occurrences:**
  - **((A B) (D (E F))**
  - **((A C) F)**
  - **(A F)**

Different begin and end points (continuation may be different)

- **Partial; have only some of the same event occurrences:**
  - **(B E)**
  - **(B D)**

Human Interaction T-Data

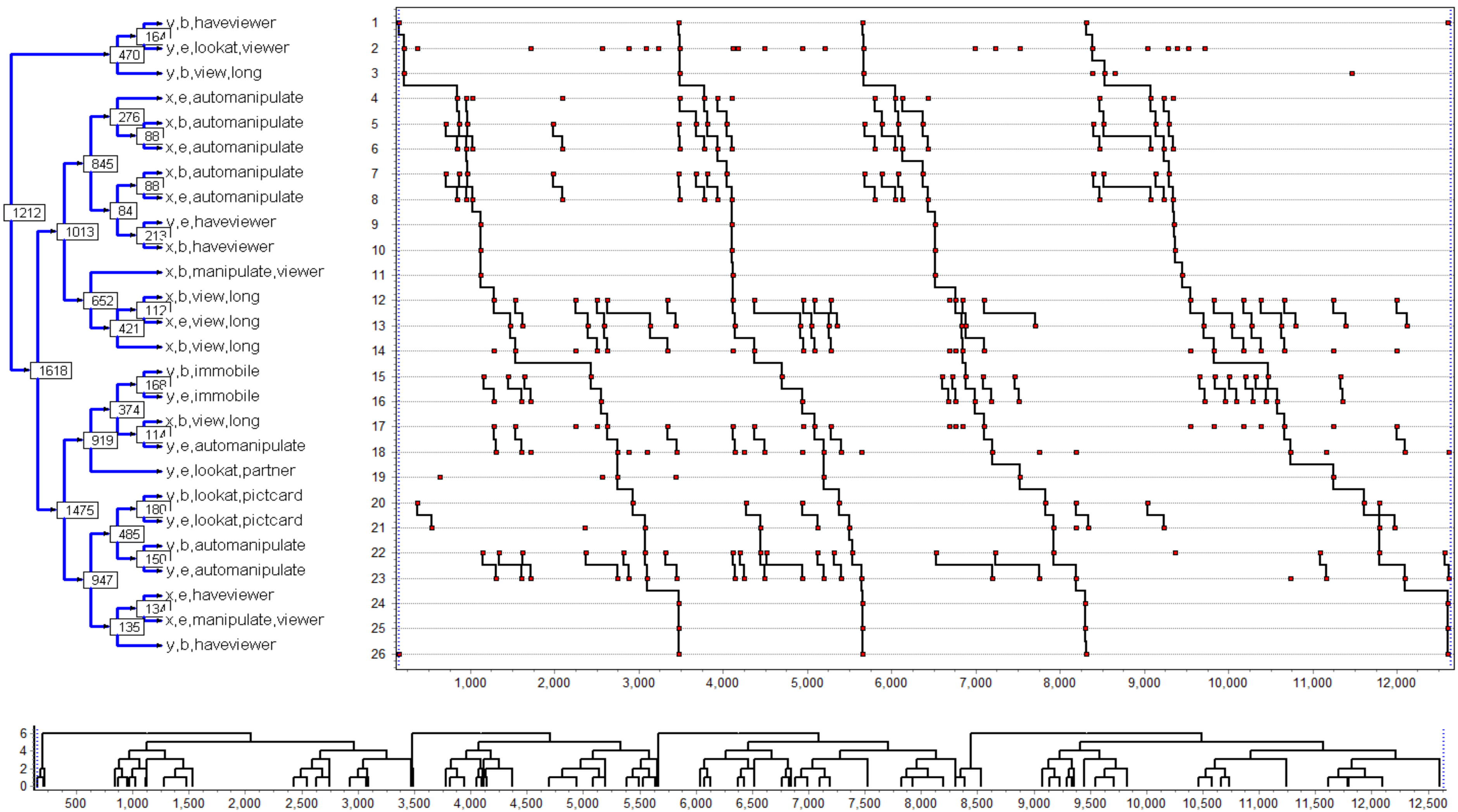
82 Event-Types, approx. 40 per child, 13:30 min



Time (one frame)1/15 s)

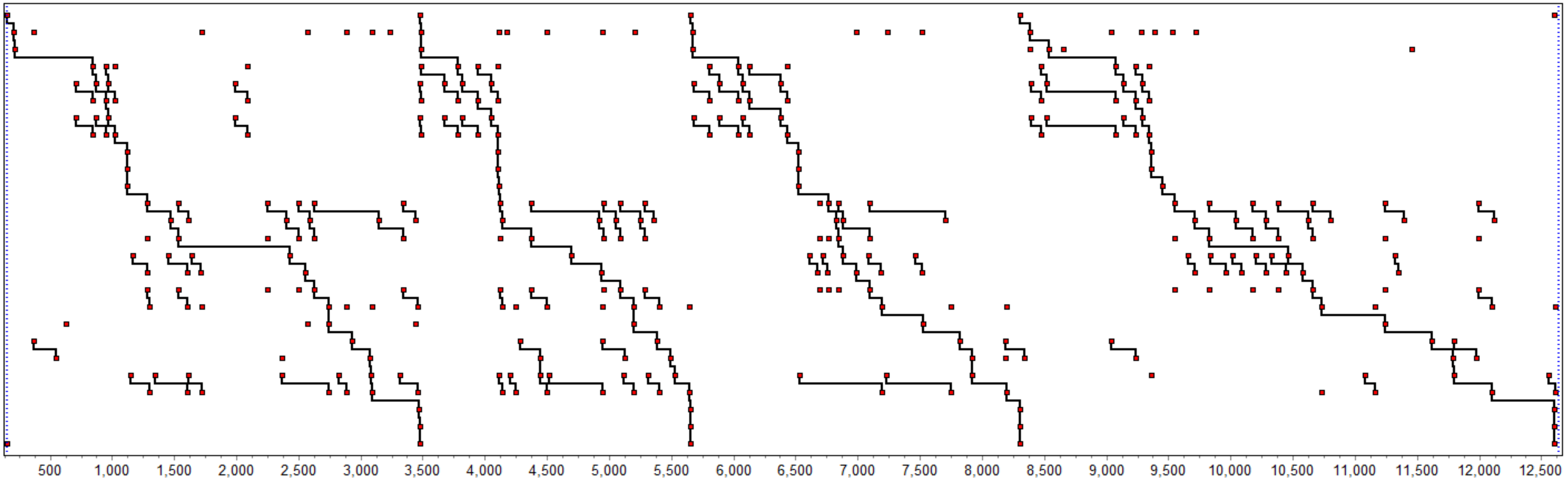
Detected T-pattern

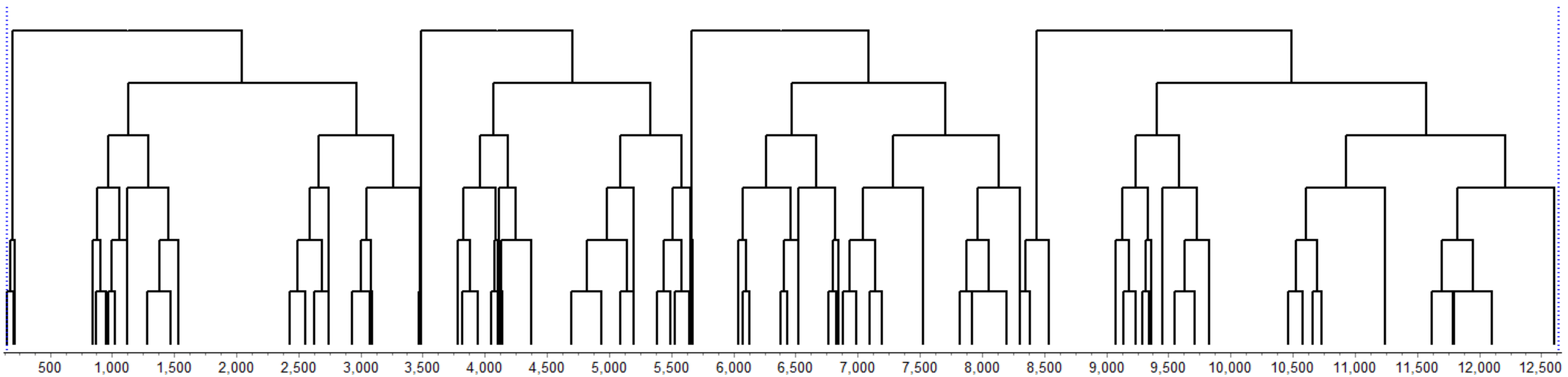
Interaction T-Pattern with 100% "Coverage"



The Same Pattern

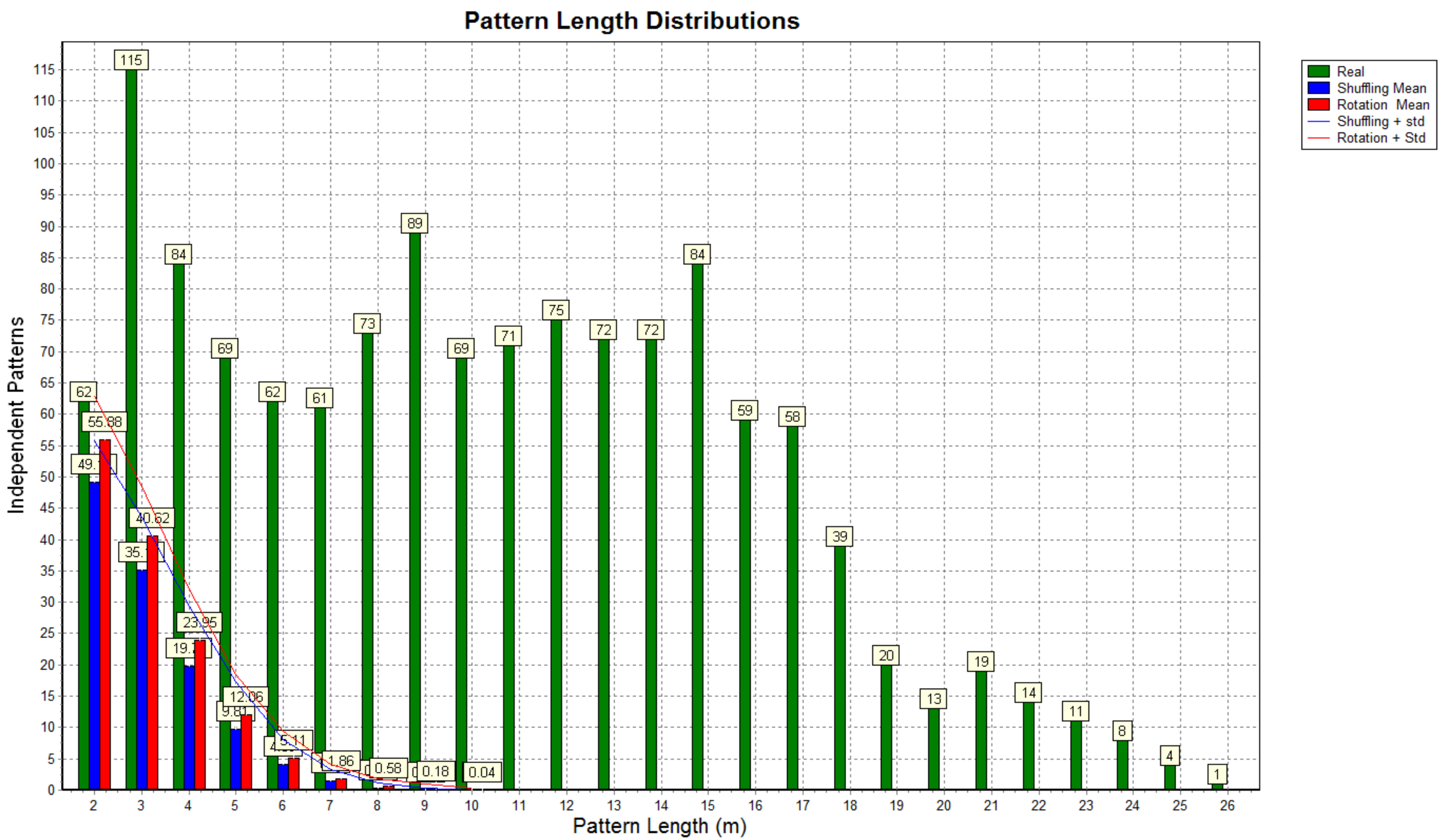
Different View





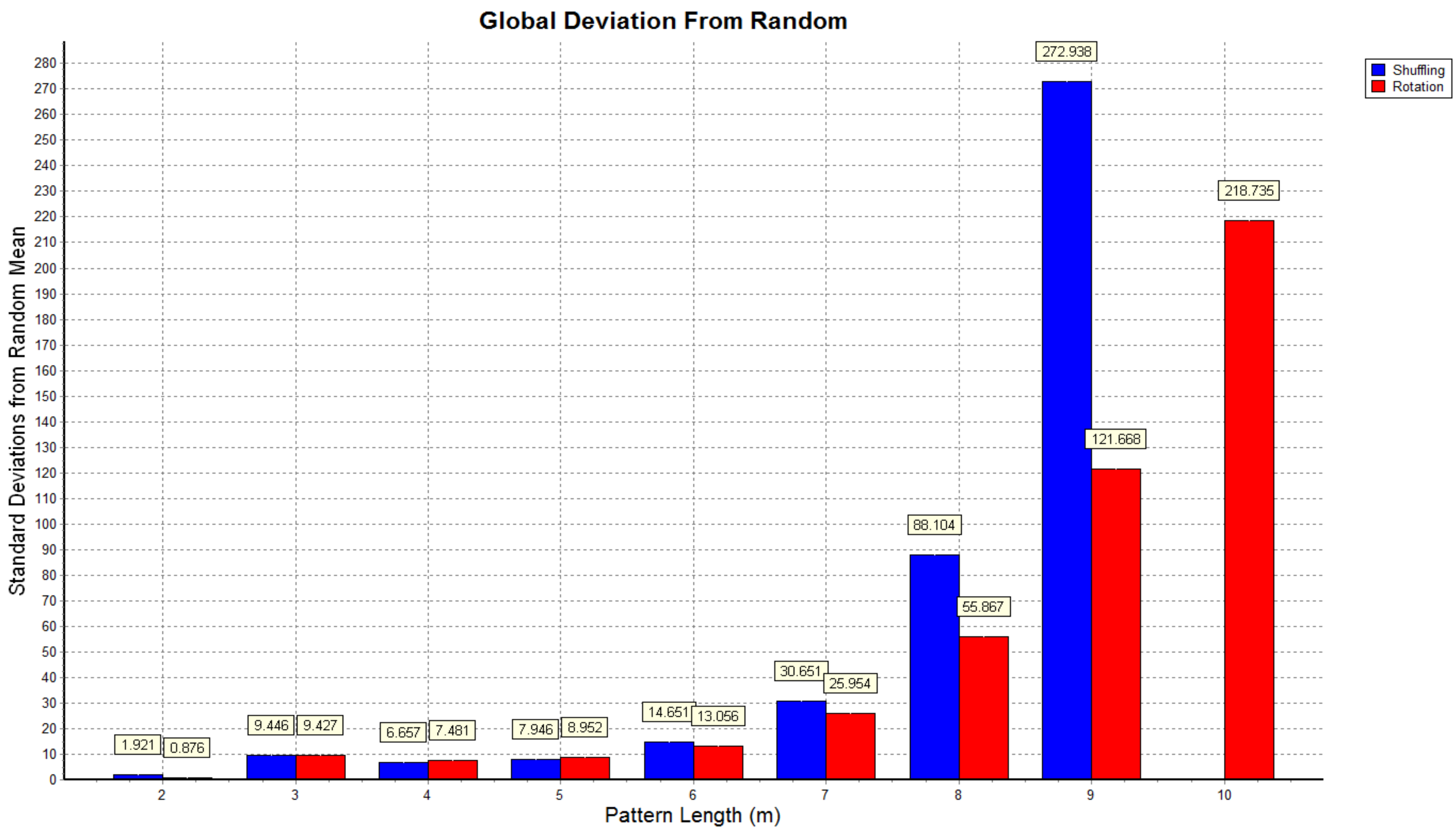
Randomized vs. Real Data (green)

100 randomizations



Randomized vs. Real Data

100 randomizations



Causal or Not Causal

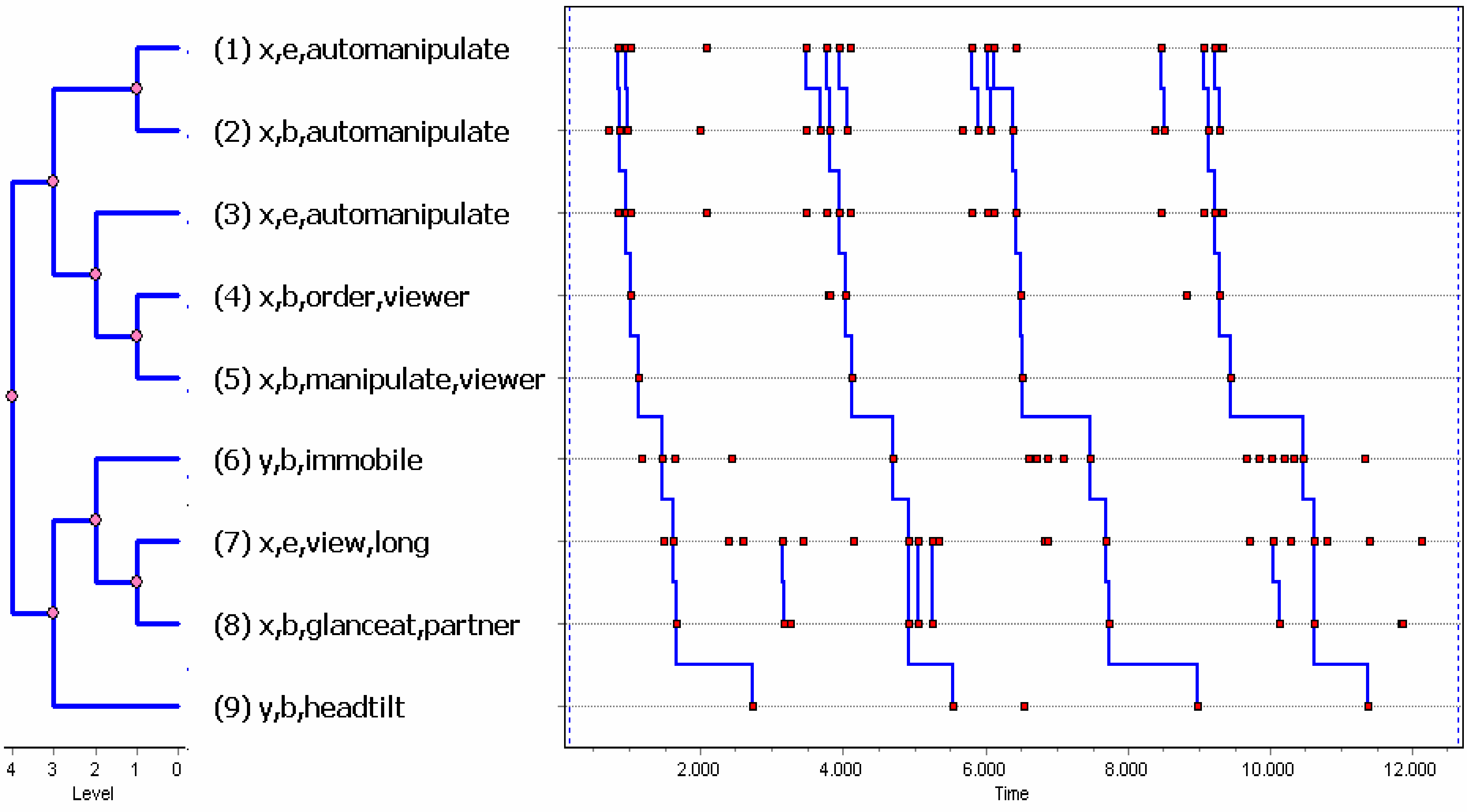
The T-patterns are the Same

- **How do you do?**
  - an earlier word is usually not considered as a cause of any word following it within such intra-individual patterns

- **How do you do? I am fine thank you**
  - an earlier part of some **inter-individual** patterns may be seen as a likely cause of a later part of the same pattern

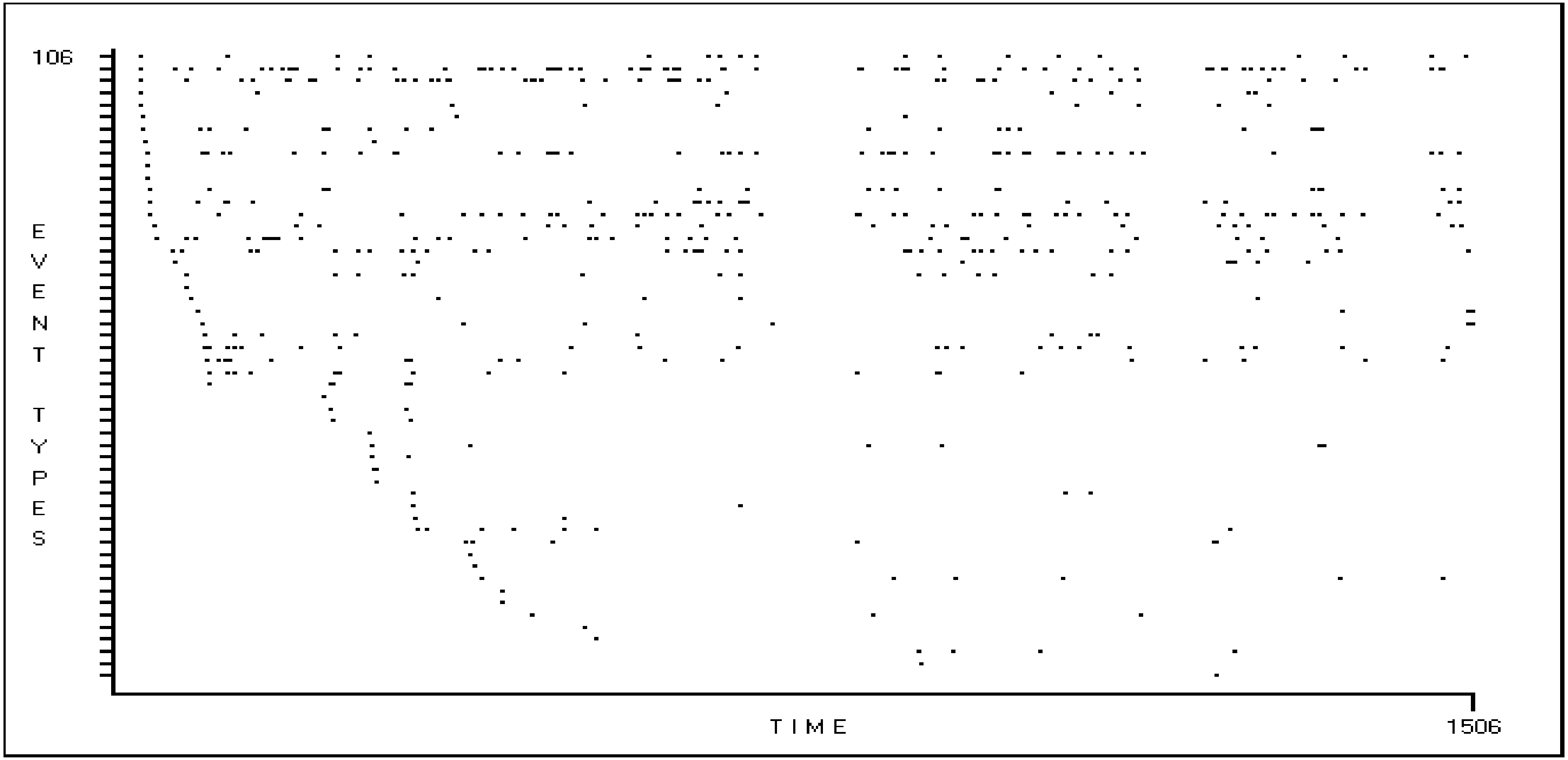
Studying Particular Behaviors

Where do Orders, Head-Tilts, and Immobile Fit In?

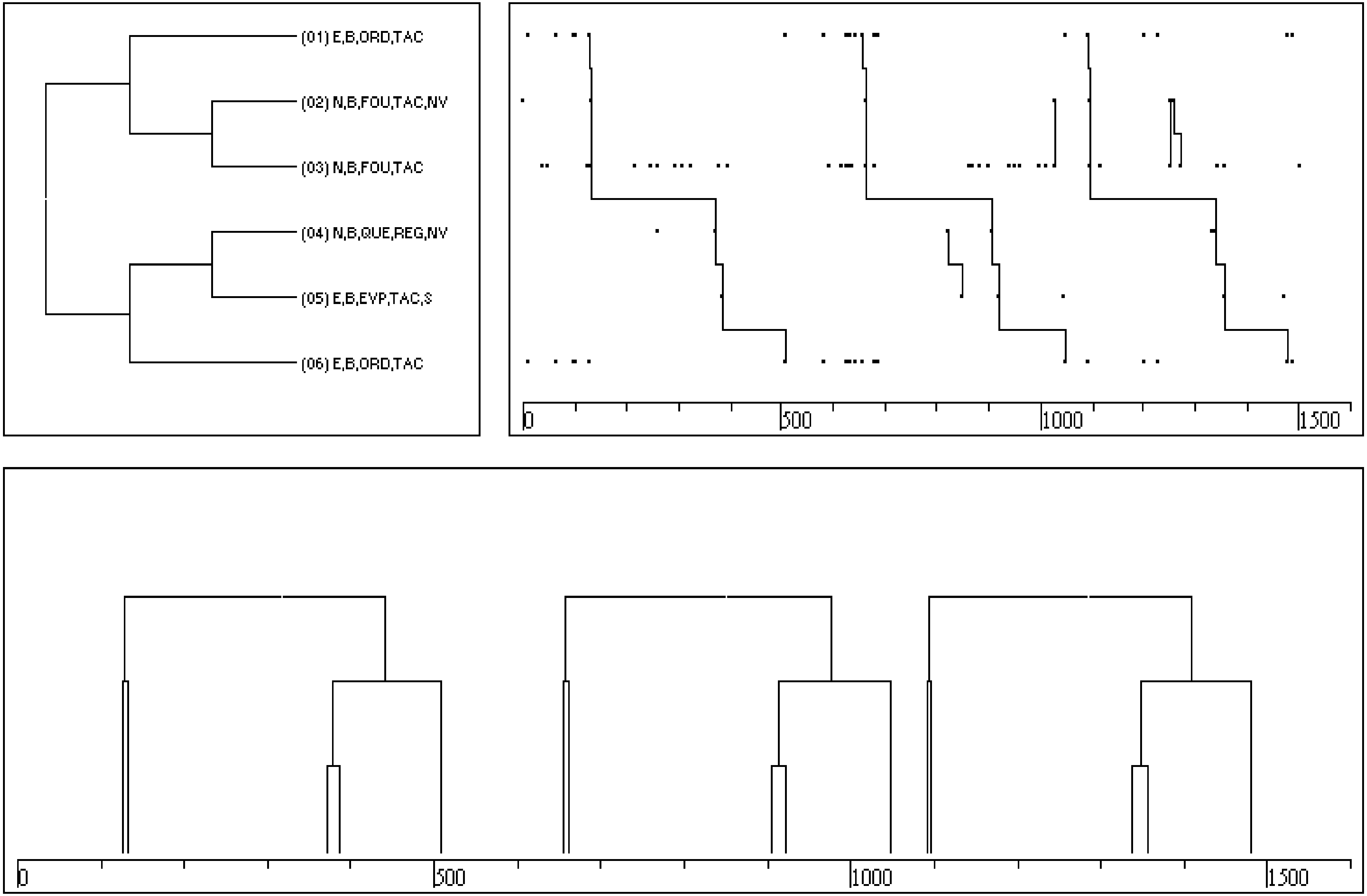


Children's Dyadic Problem Solving

25 min



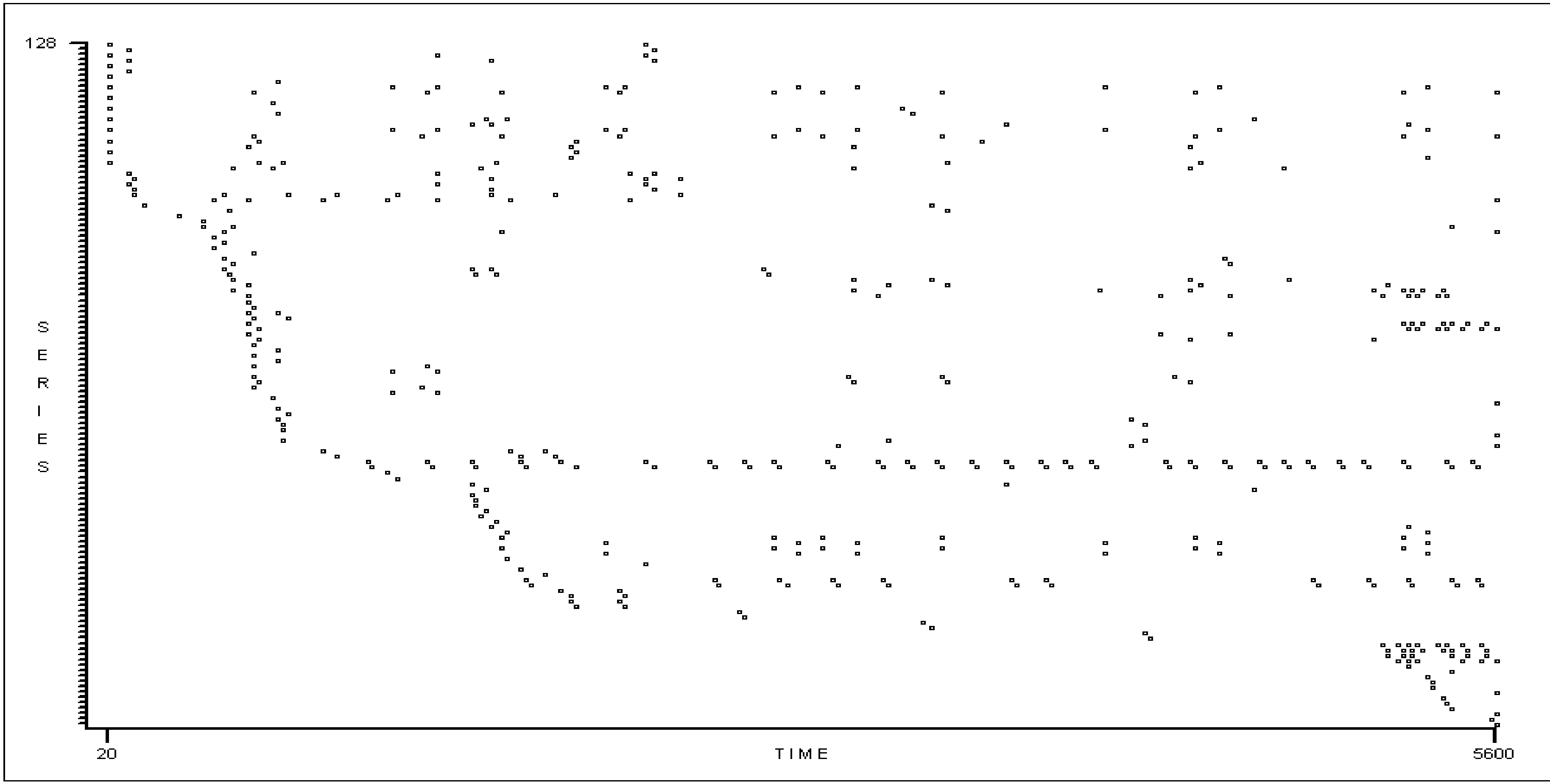
Data from published studies by Beaudichon and Magnusson.



Data from the University of Paris V. - J. Beaudichon et al.

Doctor-Patient Facial Interaction

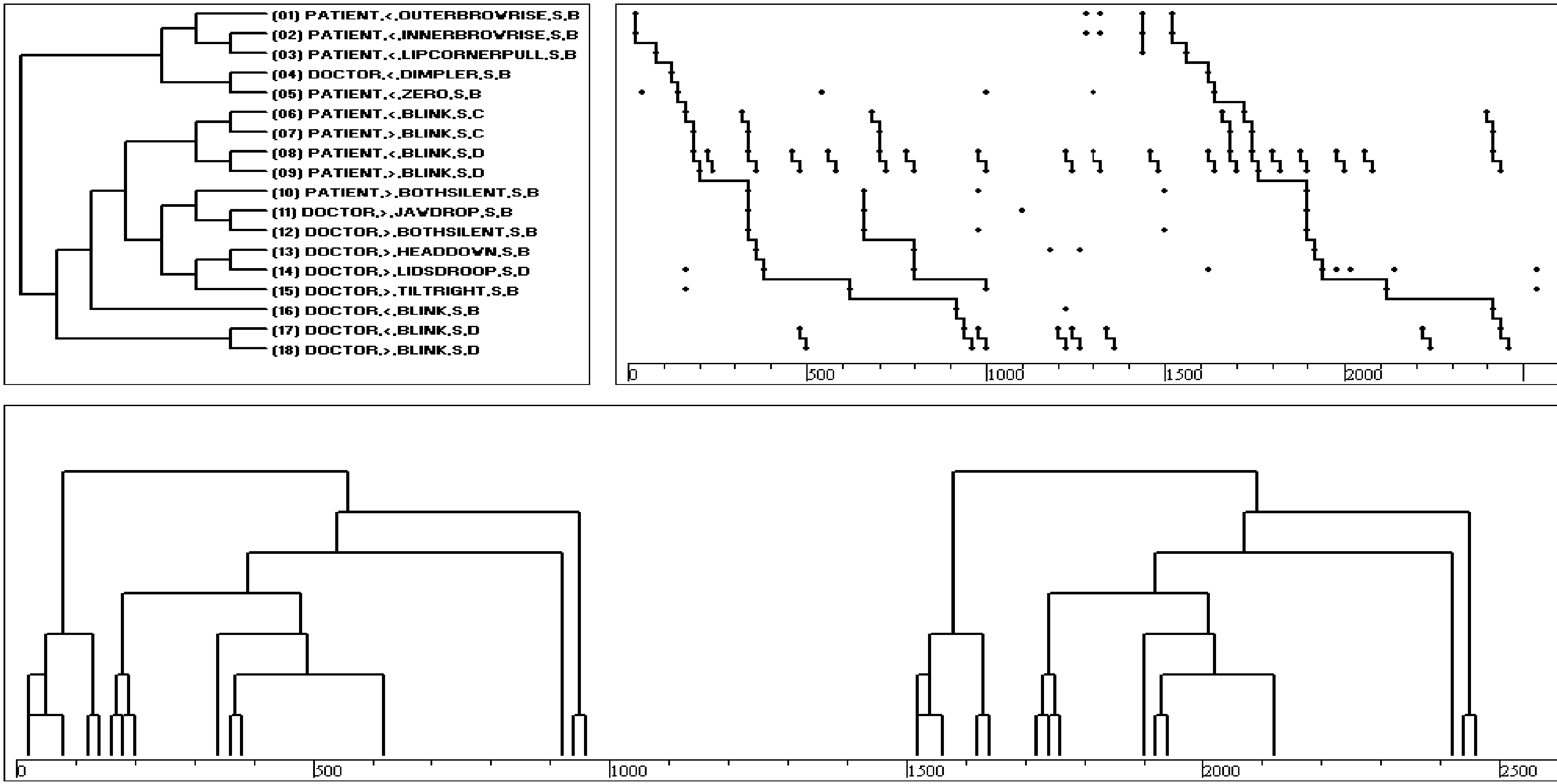
Coded with FACS, approx. 2 min



Data from the Psychiatric Hospitals Geneva, V. Haynal et al

Doctor-Patient Facial Interaction

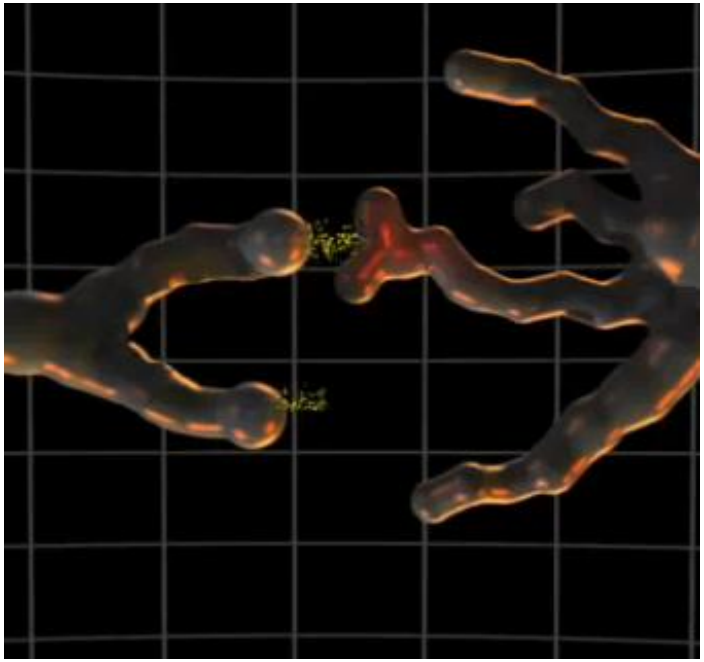
Doctor Poses Two Questions

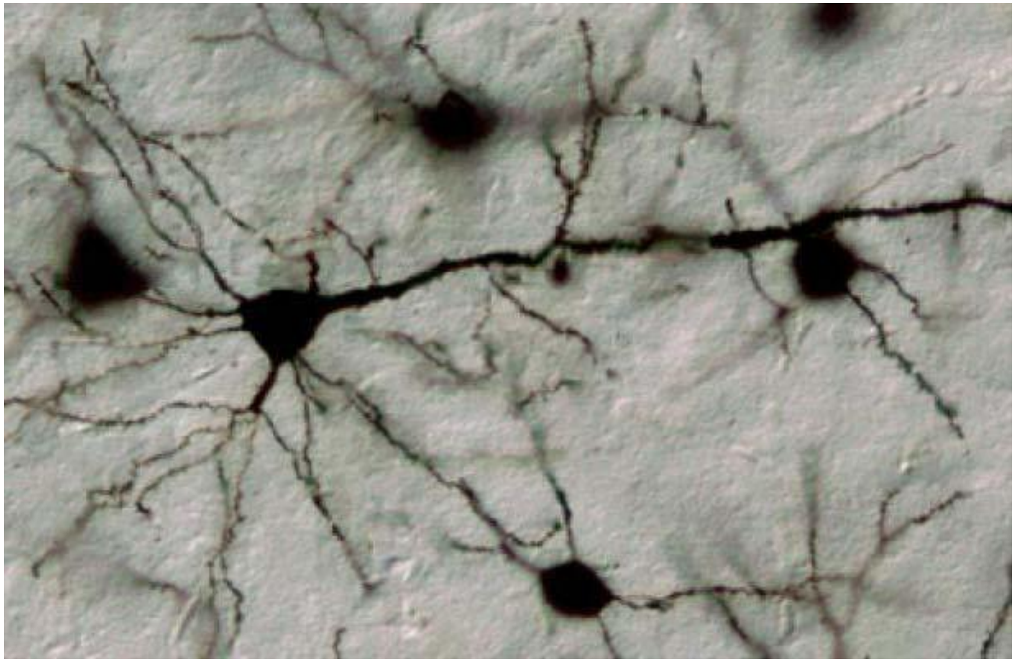


Data from the Psychiatric Hopitals in Geneva, V. Haynal et al.

Neuronal Interaction Networks

Interacting Organisms

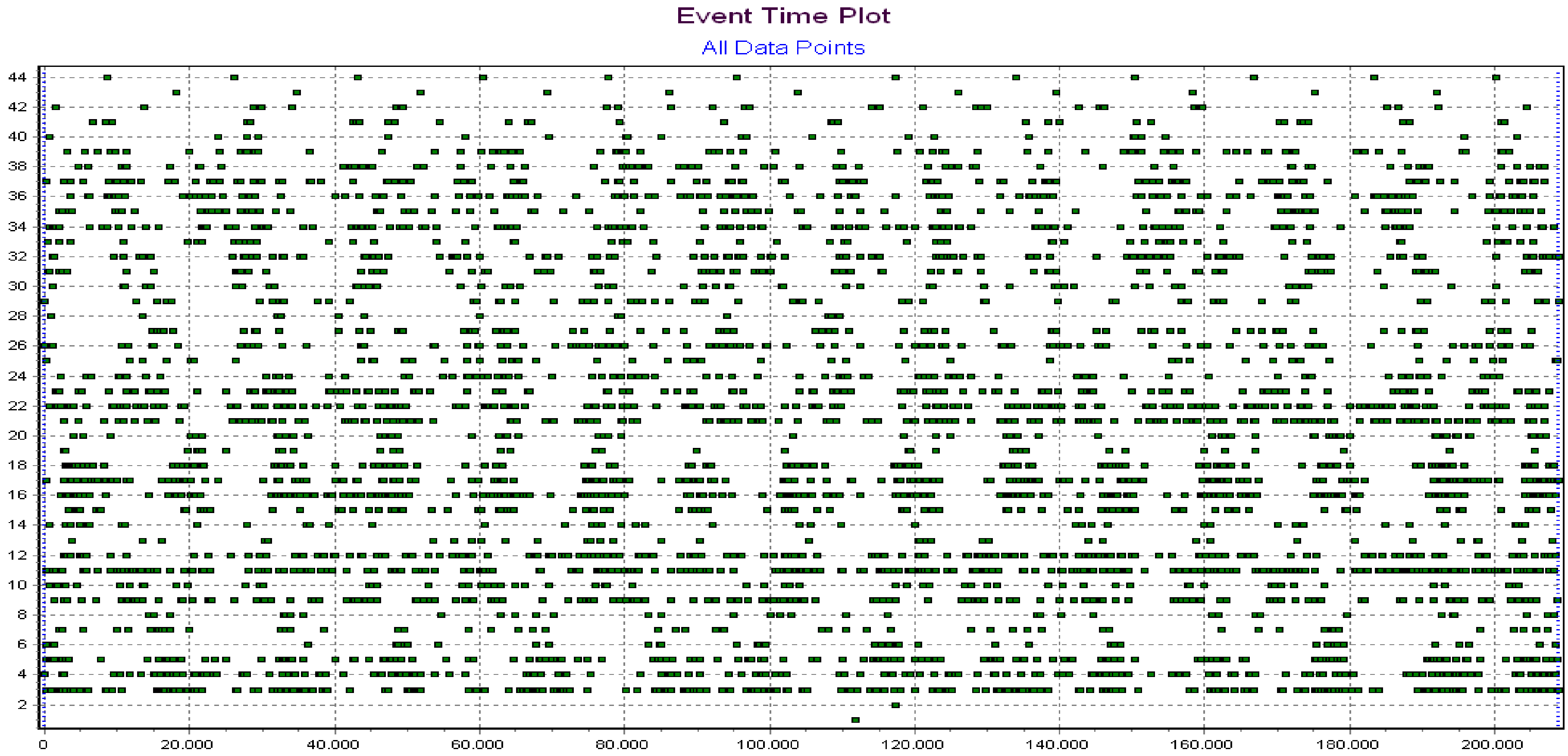




- The firing moments of tens or hundreds of individual neurons close to each other was simultaneously registered, but there is no registration of direct synaptic connections between them.

Neuronal Population Behavior

Six Seconds of Firing of 44 Neurons

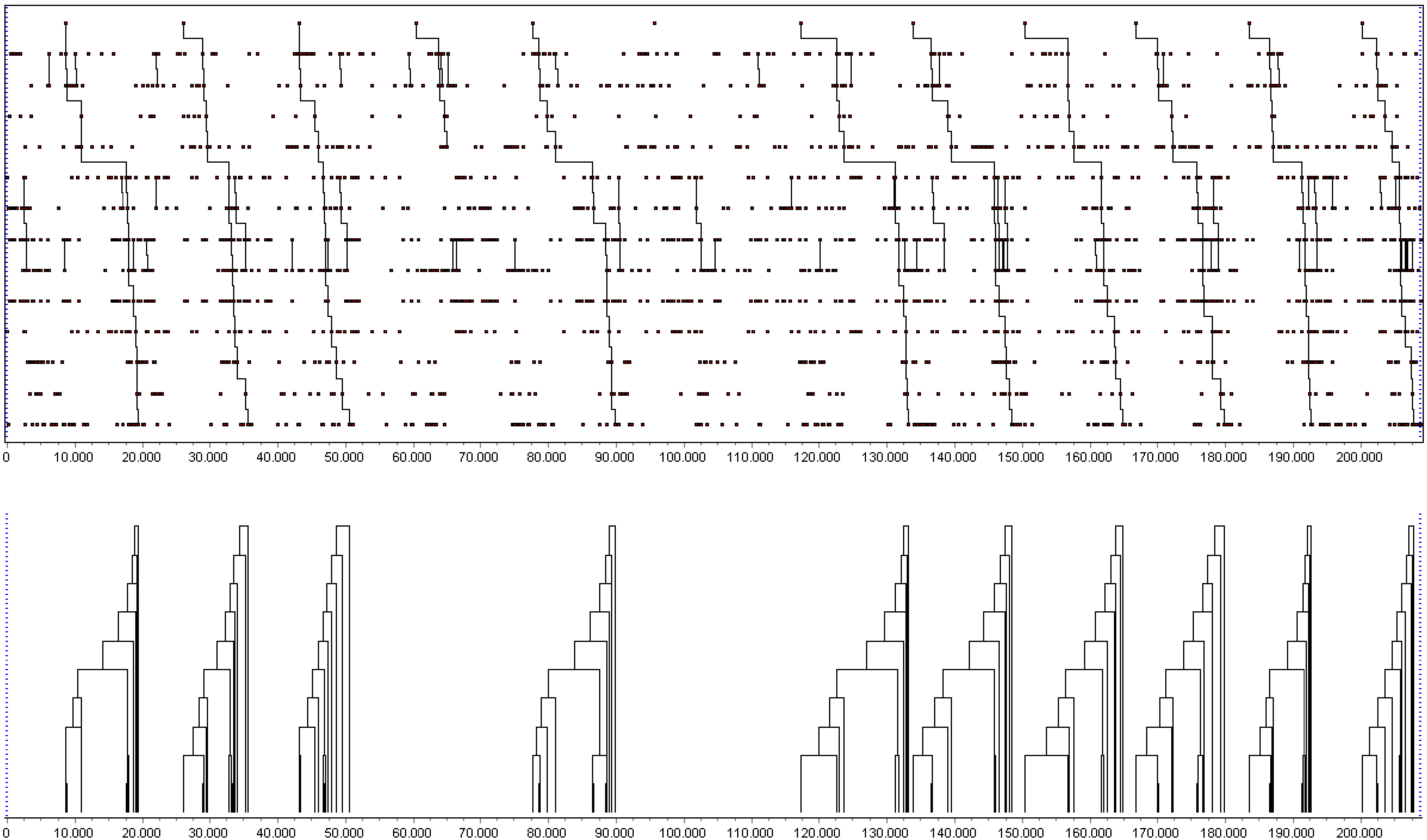


Time (3 / 1000000 s)

Registered using a microchip at the Babraham Institute, Cambridge. A. Nicol.

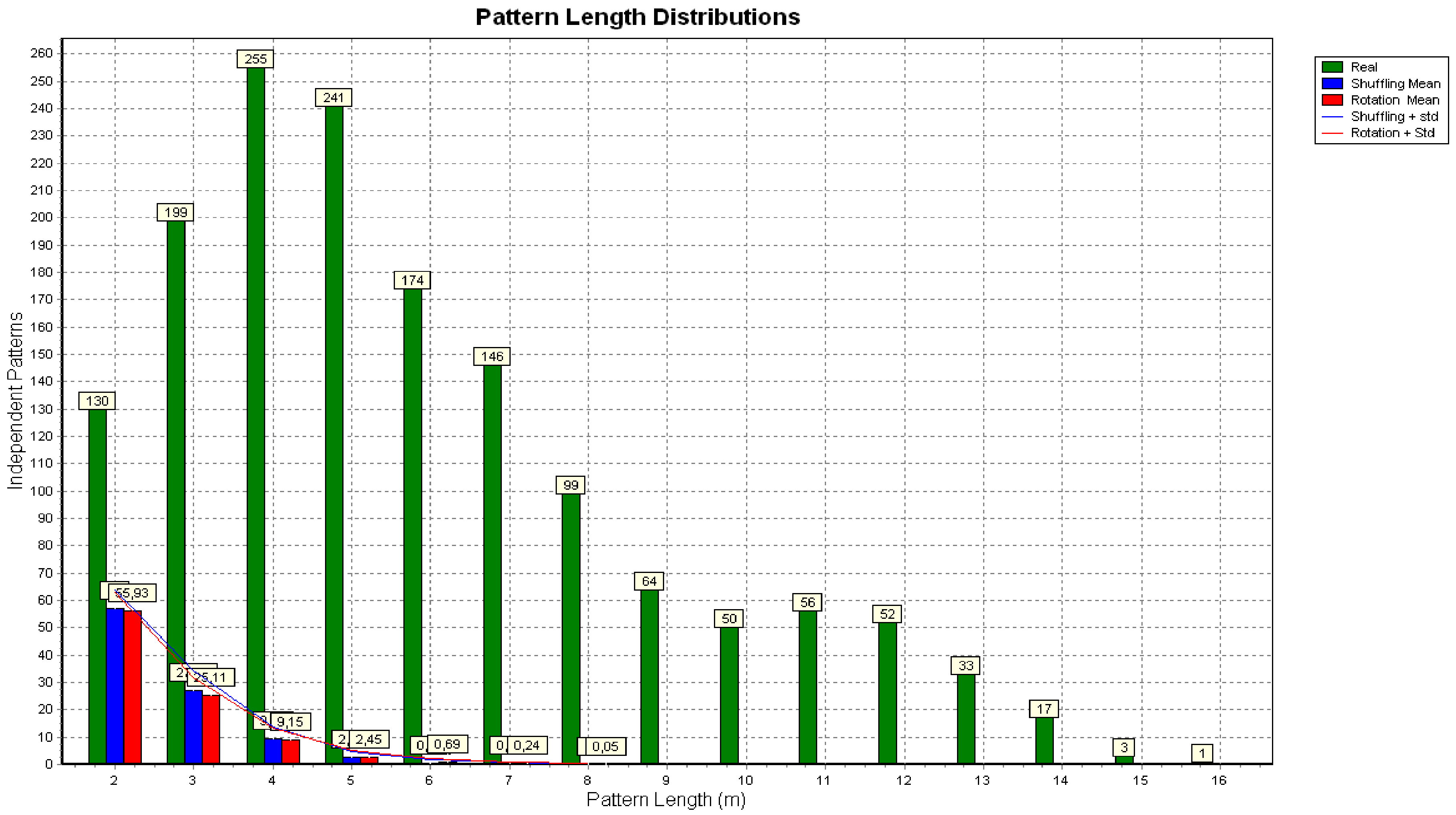
Neuronal Interaction T-patterns

12 Breathing Cycles, ≈6 seconds



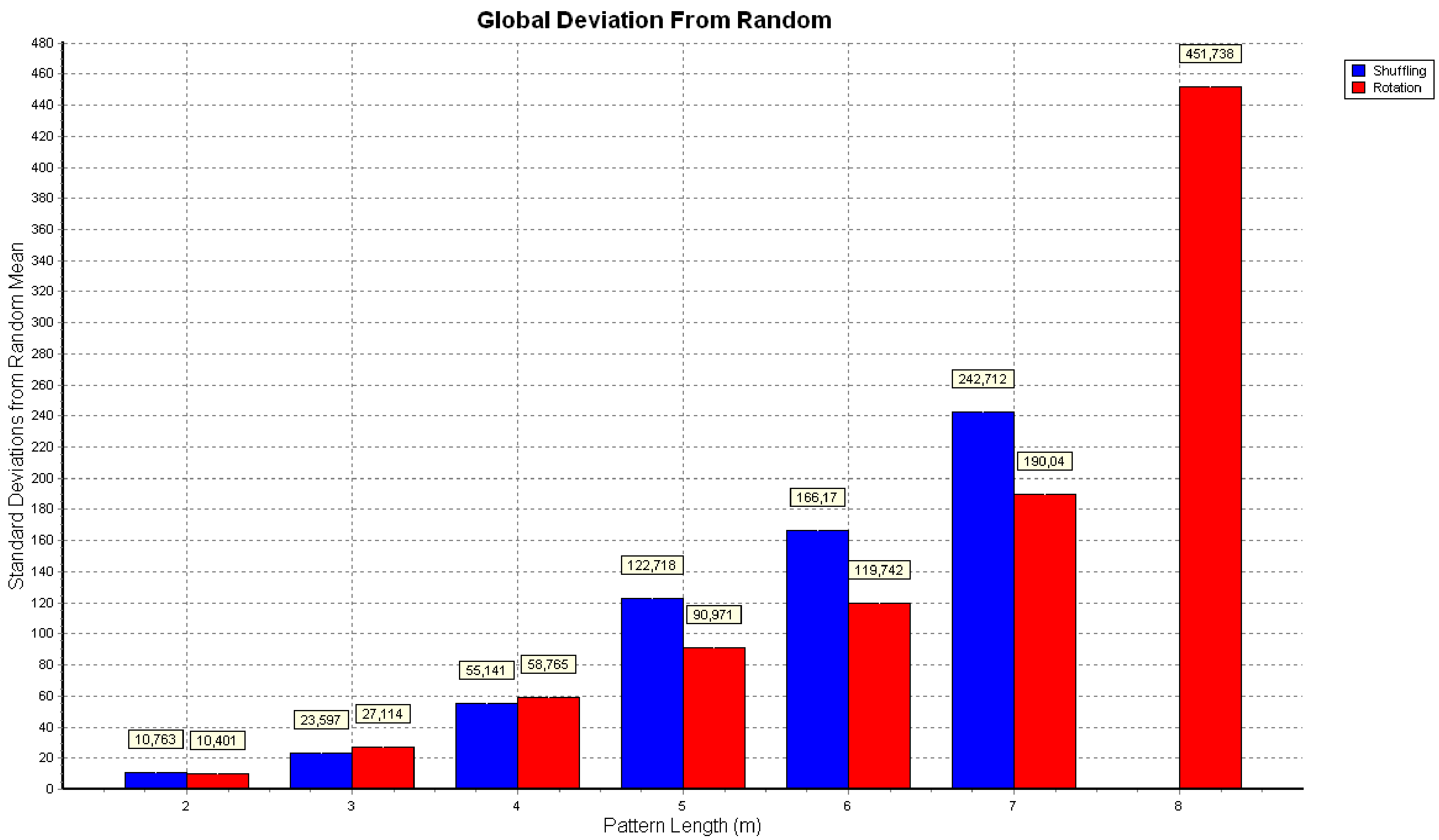
Randomized vs. Real Data (green)

100 randomizations



Randomized vs. Real Data

100 randomizations



Extending the T-model

Building on the Critical Interval and T-pattern Concept

- Markers
- Composition
- +/- Associates; Satellites & Taboos
- +/- Gravity Zones
- Packets
- Packet markers
- Drifters
- T-kappa

The K% T-marker

Indicates that a Pattern is Ongoing

- If K% of the times event-type (or T-pattern) X occurs, it occurs as a component of T-pattern Q then X is a K% T-marker of Q.

- A T-marker's occurrence thus indicates with ≥ K% probability that a particular T-pattern is occurring.

- A T-marker that occurs early in a pattern thus predicts the rest of the pattern.

- A T-marker occurring late in the pattern thus retrodicts the earlier part of the pattern.

T-Bursts

"One of the terms that now have been given definitions specially adapted to the t-system is the burst, referring to a number of points (events of the same type) occurring in succession with distances between them that are much shorter than the average. Until very recently, the t-pattern detection algorithms have not dealt directly with such phenomena which have consequently been invisible to the corresponding software. But now a "t-burst" is defined and detected as a special kind of t-pattern and can therefore also occur as a component of more complex t-patterns (including higher-order bursts). Any t-pattern can also form t-bursts, which in turn may occur as components of more complex t-patterns." (Magnusson, 2006, p.135; see link below.)

For further details as well as definitions of T-blocks and T-Cycles see:

http://dl.dropbox.com/u/1983864/Magnusson%20book%20chapter%202006.pdf

T-associates

Associates -- but Not Components -- of T-Patterns

- A positive or negative (+/-) associate of a T-pattern is: some behavior that is not a part of that pattern, but occurs within or around significantly more of its occurrences (positive) or less (negative) than expected by chance.

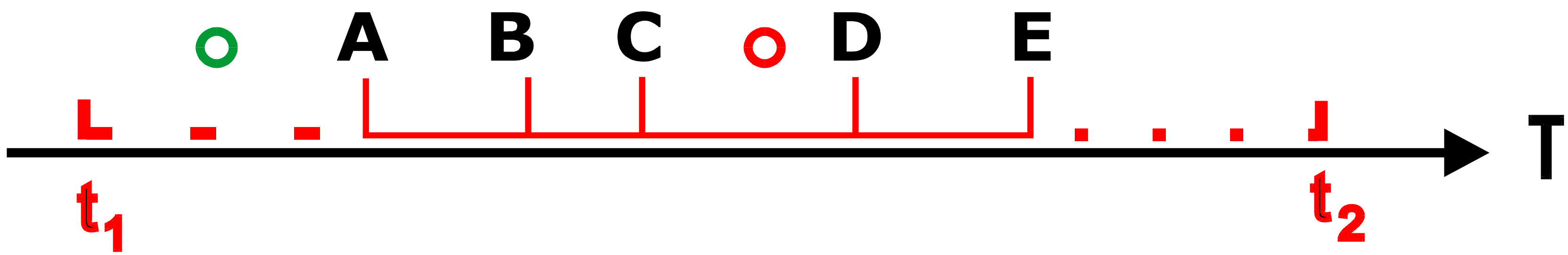
- Such associates may occur **only**, **always**, **sometimes** or **never** within or near their corresponding T-pattern.

- The "**only and always**" case is called a **T-satellite.**

- The **never** case is called a **T-taboo.**

The T-packet Structure

A T-pattern with Associates and Gravity Zones



- An instance of a **T-packet** showing two T-associate instances
- The **gravity zone**, [t_1, t_2], of a T-pattern extends from its earliest to its latest occurring positive associate.
- The **negative gravity** or **repulsion zone** (not shown) is similarly the interval within which negative associates tend **not** to occur.
- T-packets are thus simultaneously **sequential and non-sequential** repeated real-time patterns.

Application

- T-pattern detection with THEME has now been applied in numerous and diverse studies.
- New studies will be presented at the **Theme User Meeting** immediately following this tutorial.
- Many studies using THEME have been successful because of the detection of otherwise overlooked experimental effects.

Numerous examples can be found and many downloaded here:

www.hbl.hi.is/hbl_publication_references.htm

Including Anolli et al (eds.) 2005: www.vepsy.com/communication/volume7.html

Some New Features in

Theme 6

Now going through final testing before release

- 10 times faster detection.

- Detects T-bursts, positive and negative T-associates and T-Packets and offers special graphic presentations for these.

- Improved random simulation module now with an additional more conservative randomization method.

… new in Theme 6 …

- Extended pattern selection/query functions.
- Extended pattern comparison across samples.
- Many kinds of tables are easily generated allowing further statistical analysis and comparison of results obtained under different conditions.
- The data base has been completely redesigned for increased robustness and speed.

Graphical Pattern Presentation

- During the tutorial, examples of T-bursts, T-associates and T-packets are shown using new graphical features in Theme 6.

Further Information

- http://dl.dropbox.com/u/1983864/Magnusson%20book%20chapter%202006.pdf
- http://dl.dropbox.com/u/1983864/Magnusson%20book%20chapter%202005.pdf
- http://dl.dropbox.com/u/1983864/Magnusson_2000_BRMIC.pdf
- www.vepsy.com/communication/volume7.html

  www.iospress.nl/loadtop/load.php?isbn=9781586035099