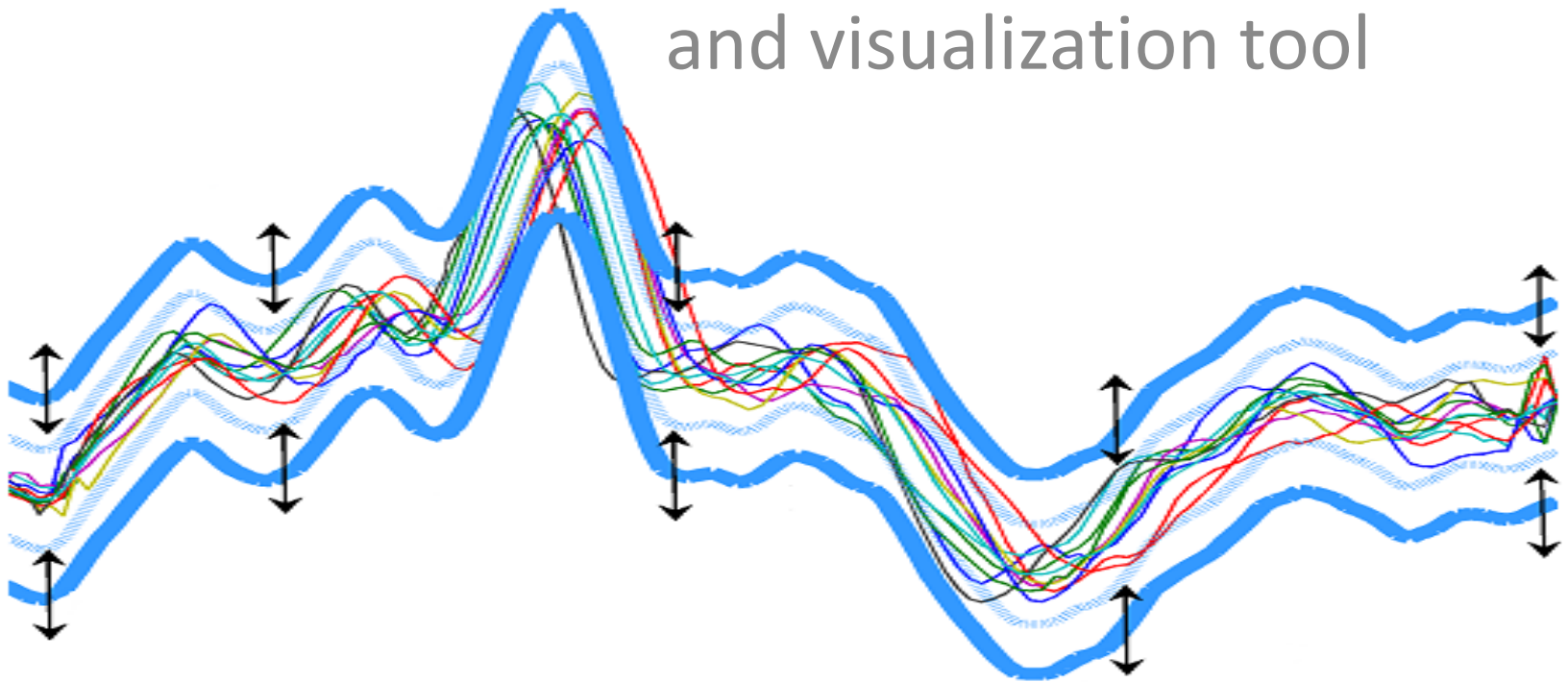


iMotifs

an interactive *multiresolution*

time series motif discovery

and visualization tool



Nuno C. Castro and Paulo J. Azevedo

Time Series Database Directory

A:\data\10



Run

 database version subsequence version

Motif Length 1000

Min Cardinality 4

Overlap% 20

Max Cardinality 32

 MinThreshold

2

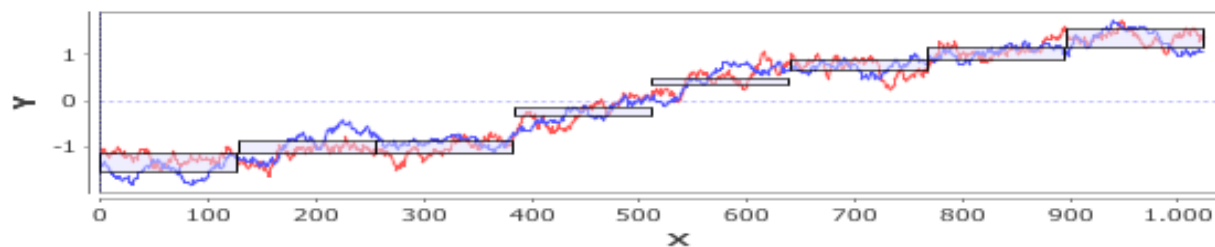
Motif Hierarchy Tree

- 0,0,0,1,2,3,3,3⁴ (163)
 - 0,1,1,3,5,6,6,7⁸ (8)
 - 1,2,2,6,10,12,13,14¹⁶ (2)
 - 1,2,2,7,11,13,12,14¹⁶ (1)
 - 1,2,3,6,11,12,12,14¹⁶ (1)
 - 1,2,2,6,11,12,12,14¹⁶ (1)
 - 1,2,3,7,11,12,13,14¹⁶ (1)
 - 1,2,2,7,11,12,13,14¹⁶ (1)
 - 1,2,2,6,10,13,13,14¹⁶ (1)
 - 0,1,1,2,4,6,6,7⁸ (6)
 - 1,2,3,4,9,12,12,15¹⁶ (1)
 - 1,3,3,5,9,13,13,14¹⁶ (1)
 - 1,2,3,5,9,12,13,14¹⁶ (1)

Invert Tree

 show all

Original Time Series

 Draw Motifs Send to back

Maximize

Finished. Idle.

Ok

Usage

- Loading the data

Time Series Database Directory

A:\data\10

Run

database version


subsequence version

Motif Length 1000

Overlap% 20

Min Cardinality 4

Max Cardinality 32



- Selecting options
- Running the MrMotif algorithm

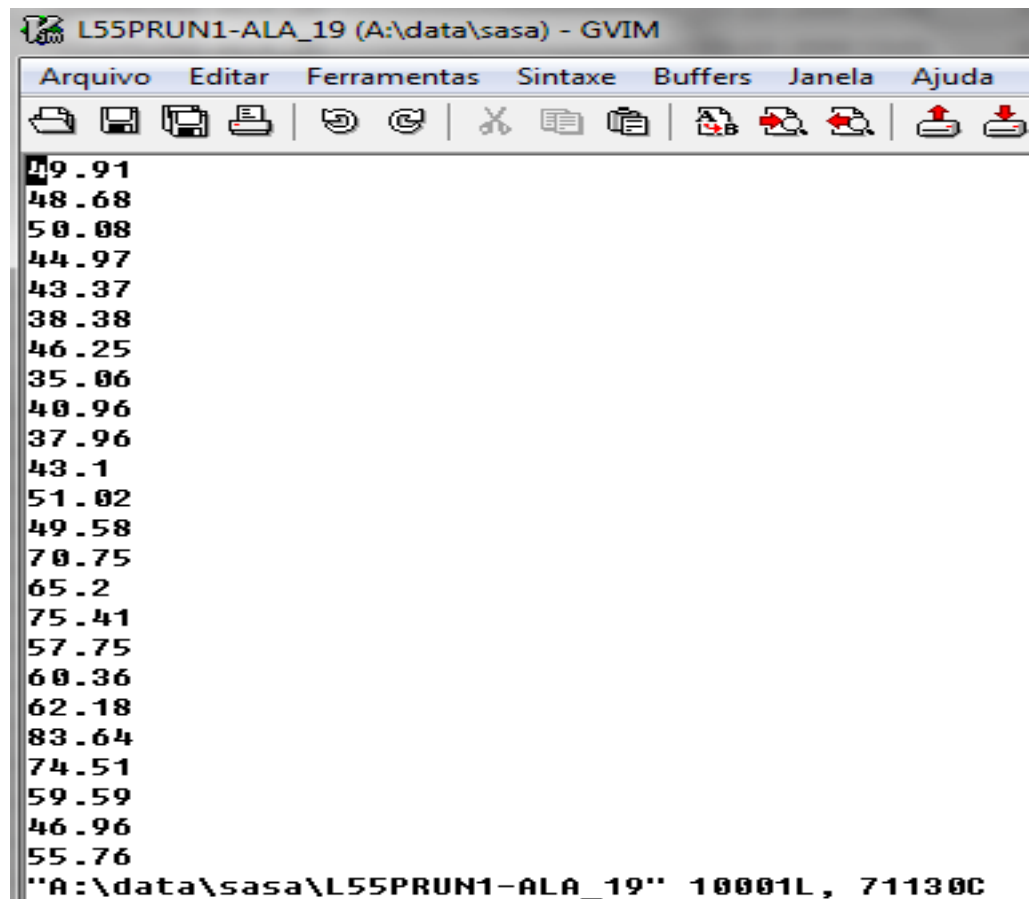
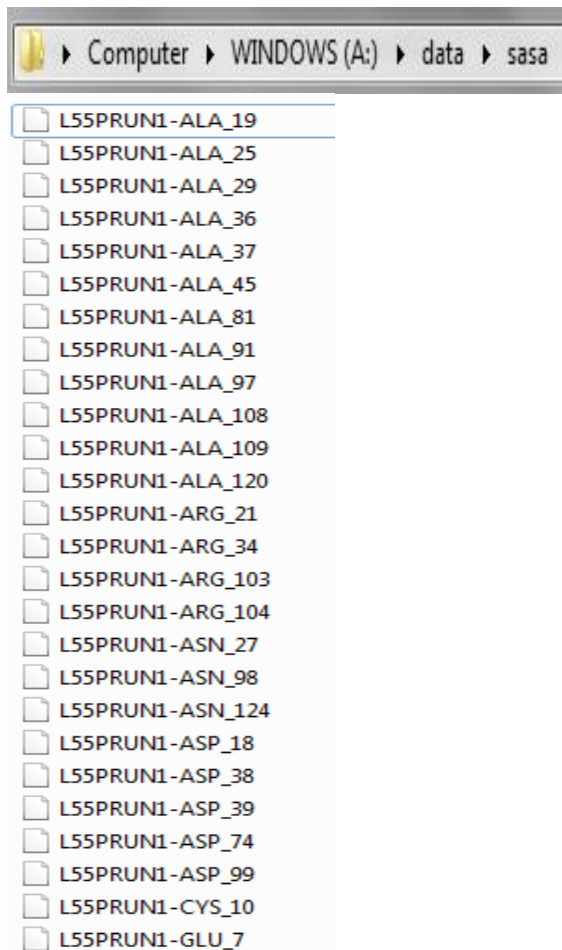
Loading the data

- Select a directory with your time series files
- Each file can be:
 1. A time series (one point per line);
 2. A set of time series (one time series per line)

This is detected automatically (no parameter necessary)
- Each time series point is separated by whitespace (not tab)

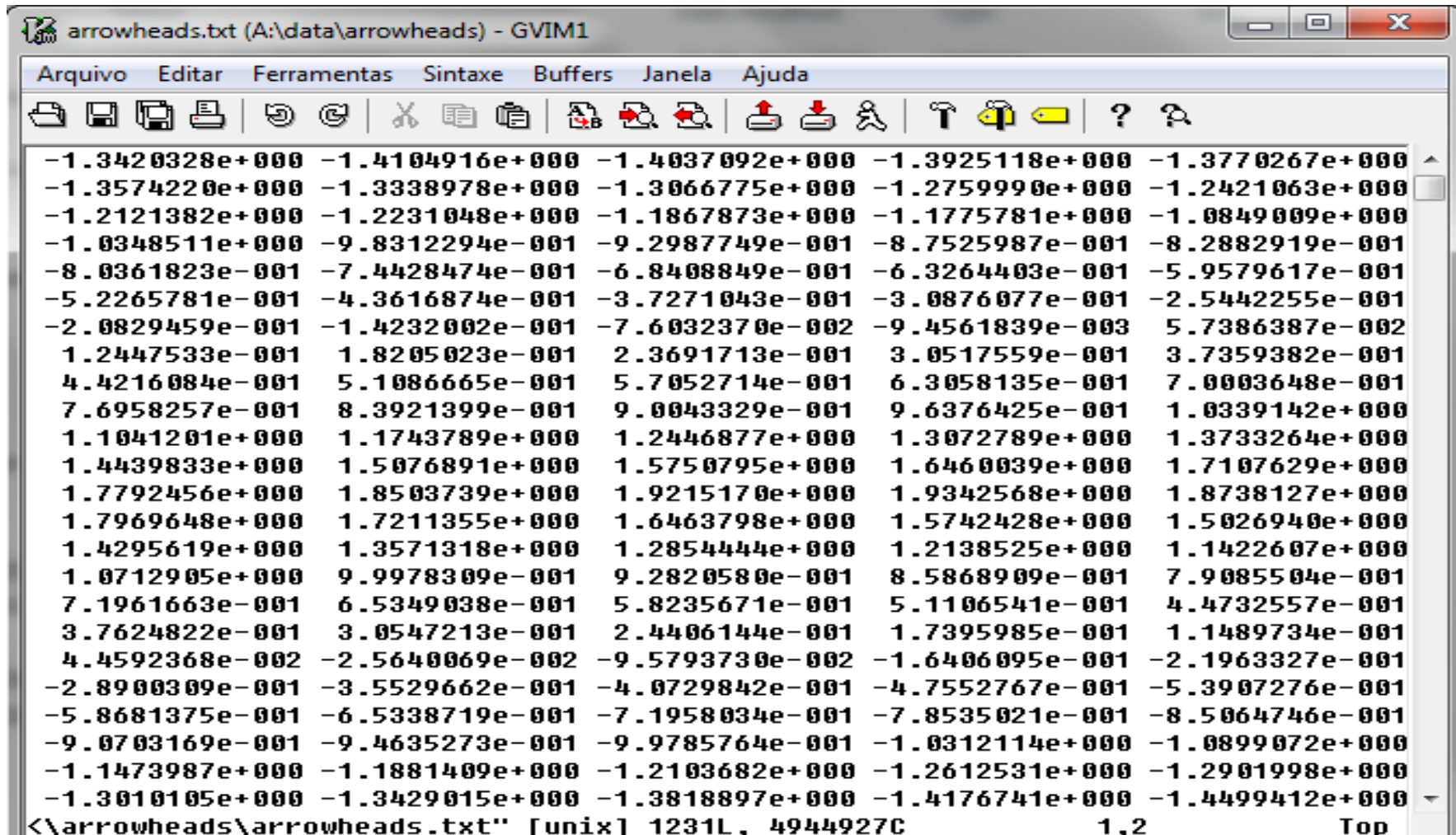
Loading the data (1)

- Example (one time series per file):



Loading the data (2)

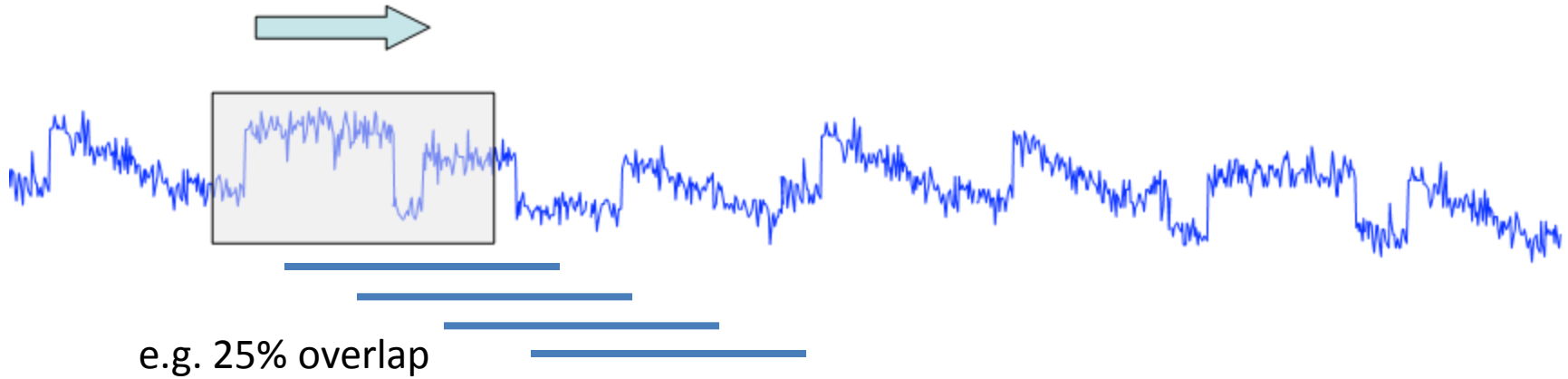
- Example (multiple time series per file, one per line):



```
arrowheads.txt (A:\data\arrowheads) - GVIM1
Arquivo  Editar  Ferramentas  Sintaxe  Buffers  Janela  Ajuda
-1.3420328e+000 -1.4104916e+000 -1.4037092e+000 -1.3925118e+000 -1.3770267e+000
-1.3574220e+000 -1.3338978e+000 -1.3066775e+000 -1.2759990e+000 -1.2421063e+000
-1.2121382e+000 -1.2231048e+000 -1.1867873e+000 -1.1775781e+000 -1.0849009e+000
-1.0348511e+000 -9.8312294e-001 -9.2987749e-001 -8.7525987e-001 -8.2882919e-001
-8.0361823e-001 -7.4428474e-001 -6.8408849e-001 -6.3264403e-001 -5.9579617e-001
-5.2265781e-001 -4.3616874e-001 -3.7271043e-001 -3.0876077e-001 -2.5442255e-001
-2.0829459e-001 -1.4232002e-001 -7.6032370e-002 -9.4561839e-003 5.7386387e-002
1.2447533e-001 1.8205023e-001 2.3691713e-001 3.0517559e-001 3.7359382e-001
4.4216084e-001 5.1086665e-001 5.7052714e-001 6.3058135e-001 7.0003648e-001
7.6958257e-001 8.3921399e-001 9.0043329e-001 9.6376425e-001 1.0339142e+000
1.1041201e+000 1.1743789e+000 1.2446877e+000 1.3072789e+000 1.3733264e+000
1.4439833e+000 1.5076891e+000 1.5750795e+000 1.6460039e+000 1.7107629e+000
1.7792456e+000 1.8503739e+000 1.9215170e+000 1.9342568e+000 1.8738127e+000
1.7969648e+000 1.7211355e+000 1.6463798e+000 1.5742428e+000 1.5026940e+000
1.4295619e+000 1.3571318e+000 1.2854444e+000 1.2138525e+000 1.1422607e+000
1.0712905e+000 9.9978309e-001 9.2820580e-001 8.5868909e-001 7.9085504e-001
7.1961663e-001 6.5349038e-001 5.8235671e-001 5.1106541e-001 4.4732557e-001
3.7624822e-001 3.0547213e-001 2.4406144e-001 1.7395985e-001 1.1489734e-001
4.4592368e-002 -2.5640069e-002 -9.5793730e-002 -1.6406095e-001 -2.1963327e-001
-2.8900309e-001 -3.5529662e-001 -4.0729842e-001 -4.7552767e-001 -5.3907276e-001
-5.8681375e-001 -6.5338719e-001 -7.1958034e-001 -7.8535021e-001 -8.5064746e-001
-9.0703169e-001 -9.4635273e-001 -9.9785764e-001 -1.0312114e+000 -1.0899072e+000
-1.1473987e+000 -1.1881409e+000 -1.2103682e+000 -1.2612531e+000 -1.2901998e+000
-1.3010105e+000 -1.3429015e+000 -1.3818897e+000 -1.4176741e+000 -1.4499412e+000
<\arrowheads\arrowheads.txt" [unix] 1231L, 4944927C 1,2 Top
```

Options

- Subsequence Version
 - Motif are extracted using a sliding window of length *Motif Length*.
 - Each sliding window overlaps *overlap%* with the previous one (*overlap%* helps controlling trivial matches)



Options (cont.)

- Database version
 - There is no sliding window
 - We have a time series database, where each time series is a different object in the DB, e.g. Database of shape time series, where each time series represents a shape. Motifs are similar frequent shapes.

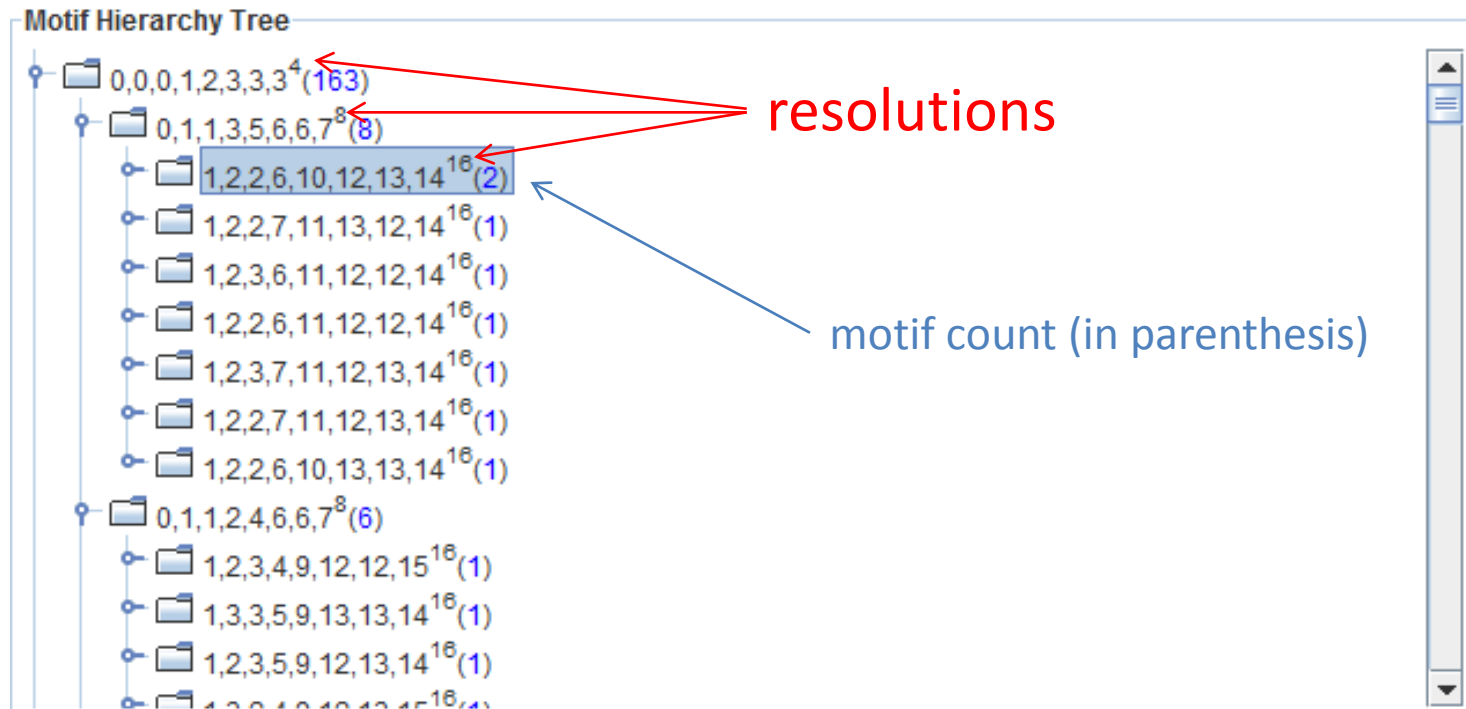
Options (cont.)

- *Minimum* and *maximum* resolution (AKA alphabet size or cardinality) can be set to 4 and 32, unless you want to play with the algorithm
- MrMotif finds motifs for all powers of 2 between *min* and *max*, e.g. 4,8,16,32
- The larger the resolution, more difficult it is to find motifs. So, motifs that persist as resolution increases are best.

→ Main idea behind the MrMotif algorithm

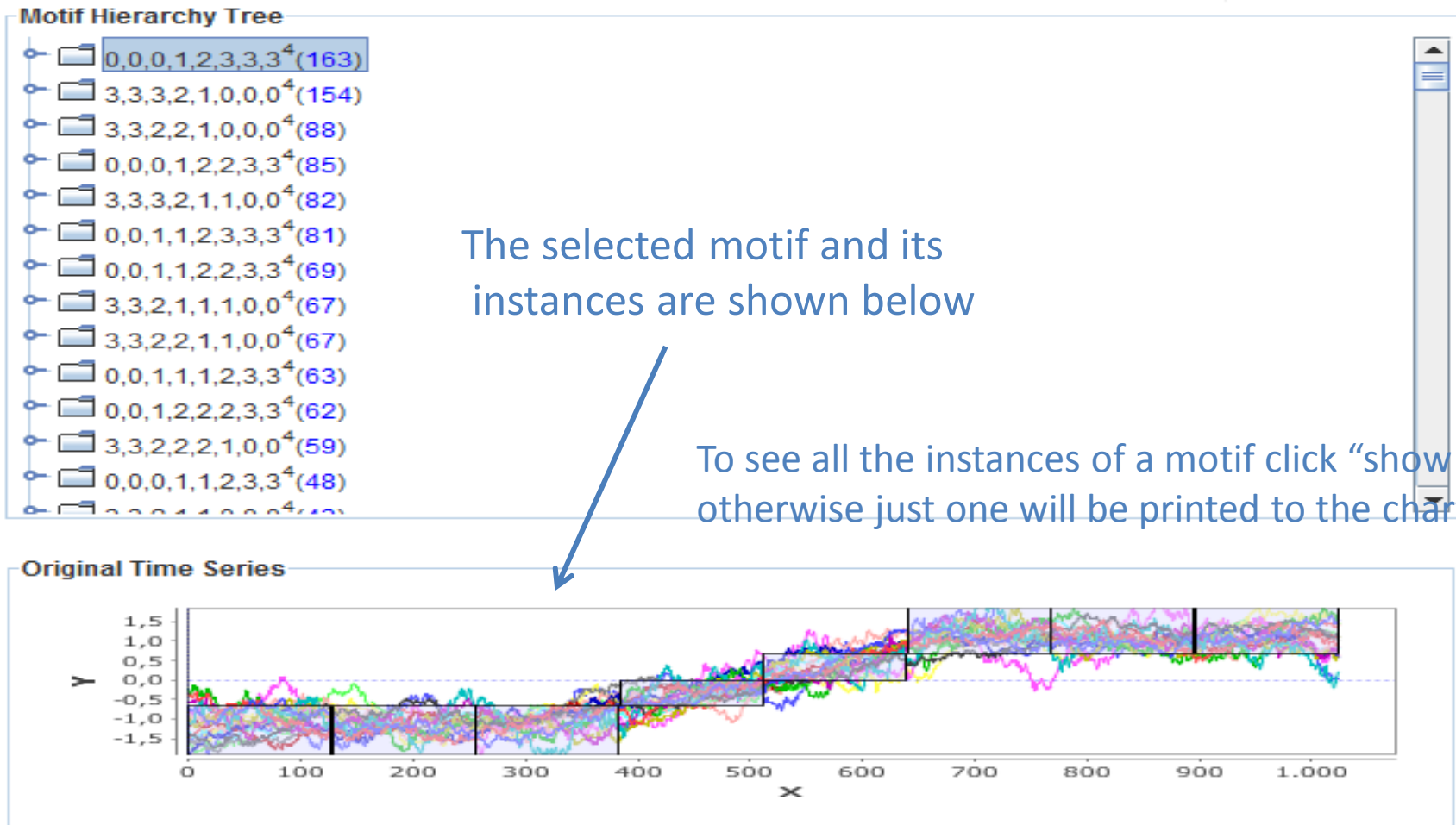
Running MrMotif and visualizing

- As MrMotif searches the DB for motifs in all resolutions between *min* and *max*, we can draw a motif hierarchy tree for each motif:



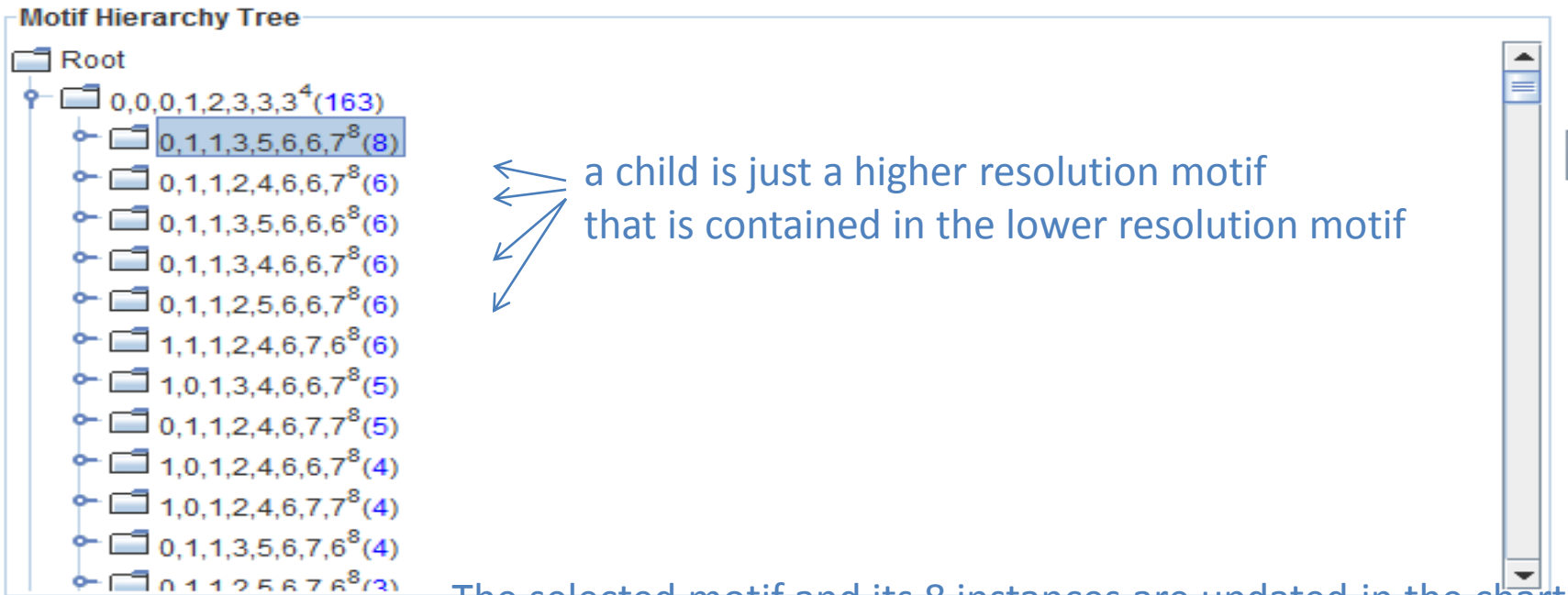
Motif Hierarchy

- First, all motifs are shown for *min* resolution (4)

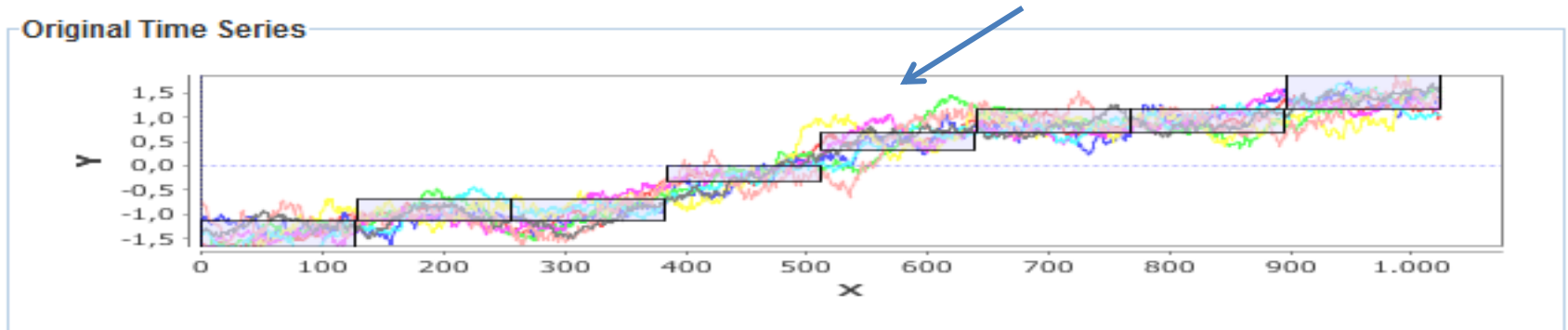


Motif Hierarchy

- By clicking a motif, we “drill-down” to its children in the next resolution (8)

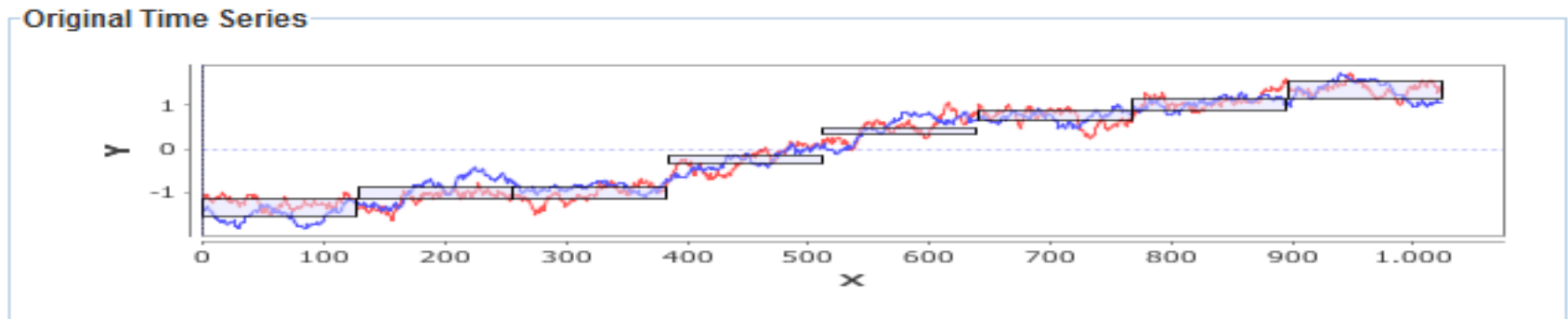
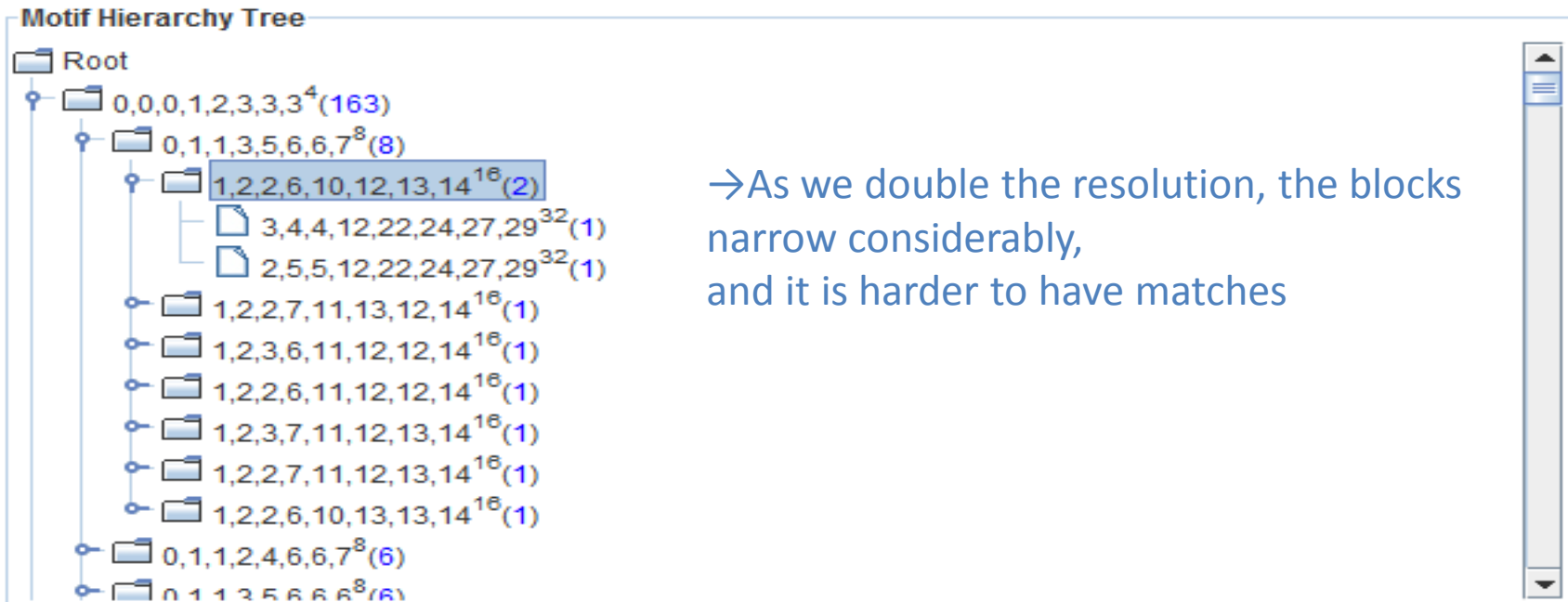


The selected motif and its 8 instances are updated in the chart



Motif Hierarchy

- And so forth:



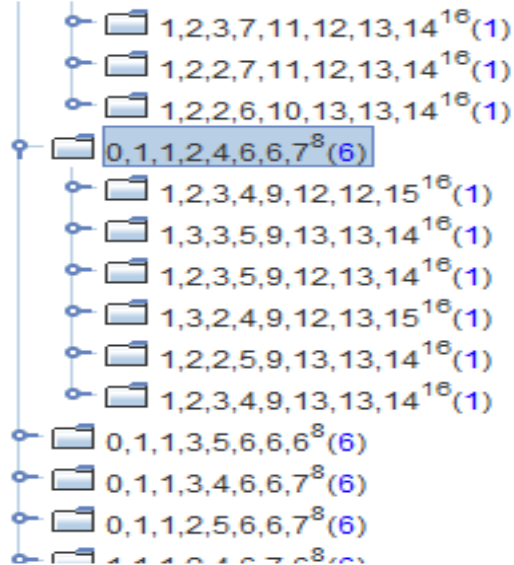
Multiresolution

- At the larger resolutions (16 and 32) it is considerably harder to have a match:
 - The time series need to be really similar to match those tiny blocks
 - Therefore we can skip expensive distance (e.g. Euclidean) calculations when we do have a match
- This is the main intuition behind the MrMotif algorithm
 - it allows the algorithm to execute faster

Example

- Take for example, motif $\{0,1,1,2,4,6,6,7\}$ at resolution 8, having 6 repetitions in the database:

Motif Hierarchy Tree

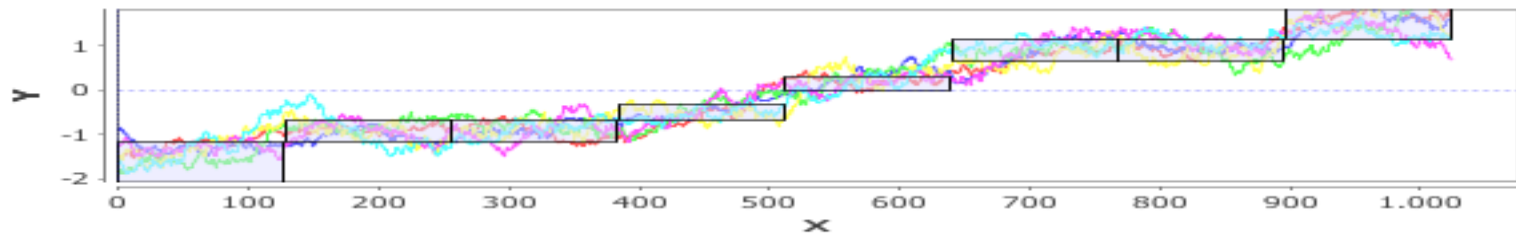


- As we increase to granularity 16, none of its children could match.

- They are not similar enough to match at the higher resolution (16).

→ Therefore, this motif can be ignored for further study.

Original Time Series



Multiresolution

- This permits the user to gain useful insight over its time series database
 - He can study how the most frequent motifs behave over the several resolutions
 - He can play with the motif hierarchy and find interesting patterns that he wasn't expecting

Other features

- Invert tree
- Location and Euclidean distance calculation

Invert tree feature

- This feature allows to invert the hierarchy tree, i.e. Show the most frequent motifs at the highest resolution (32).
- When using this feature, the *max* resolution should be set to 16 or 8, since it may be hard to find motifs at resolution 32.

Location and Euclidean dist.

- By clicking with the second mouse button on a motif in the tree, there is the possibility to:
 - Get the location of the motif instances in disk
 - Calculate the Euclidean distance among all the instances that compose a motif.

