**DEFINITION**

A time series **Motif** is a frequent pattern in time series data, i.e. a repetition of a particular subsection of the series.

**Focus**

We are particularly interested in:

1. Motif Discovery Algorithms
2. Motif Evaluation Measures
3. Motifs as Building-Blocks for other Data Mining tasks

**1 – MOTIF DISCOVERY**

We introduce Multiresolution Motif Discovery algorithm – MrMotif, a scalable algorithm to discover motifs in time series at several resolutions.

MrMotif is:

- **Fast**
  - Linear algorithm
  - One sequential disk scan
  - Constant access time structures (hashtables)

- **Space-Efficient**
  - Uses little memory
  - Adjustable memory
  - Space-saving algorithms

- **Intuitive**
  - Straightforward output
  - Returns most frequent patterns in the database

- **Robust to Noise**
  - Maintains the quality of patterns in noisy data

- **Easy to use**
  - A small number of parameters needs to be configured

- **Simple**
  - Algorithm is simple to understand and implement

- **Reproducible**
  - We provide the source code and datasets

**Motivation:**

- Large number of motifs returned by the algorithms
- Evaluation is subjective
- Relatively unexplored area

**Approach:**

- Information measures
- Accuracy related measures
- Statistical tests

**How to obtain a ground truth?**

- Markov Models
- Randomly generated time series

- **ongoing work**

**2 – MOTIF EVALUATION**

**Motivation:**

- Time series motifs are meaningful patterns which are characteristic of a particular application domain

- As motifs can be used to describe the time series, they can be used as “building blocks” for other data mining tasks:
  - Classification
  - Abnormality detection
  - Forecasting

**ongoing work**

**3 – MOTIFS AS BUILDING-BLOCKS**

**Motivation:**

- Markov Models
- Randomly generated time series
- Future work

*Contact: castro@di.uminho.pt*

Web site: www.di.uminho.pt/~castro/mrmotif