# Software Tool for Cardiologic Data Processing

Michal Huptych, Lenka Lhotská  
Czech Technical University in Prague, FEE, Department of Cybernetics  
huptycm@fel.cvut.cz

## 1 Introduction

Standard ECG, multi-leads ECG (Body Surface Potential Mapping - BSPM) and intracardial signals from catheters are three types of measurement of electrical heart signals. These measurements represent complex electrophysiology investigation particularly useful in the case of experimental measurement. Along with the data about the state of a measured subject (age, anamnesis, medication, etc), we obtain very large database for complex analysis of heart function. This paper presents solution of the first step of heterogeneous data processing – design of basic data model and implementation of an application (on the designed data model) for integration, pre-processing, analysis and visualization of data. Data in the data model can be consequently used for analysis of functional relations of acquired data. This project has been solved by an interdisciplinary team of experts in biomedical engineering and medicine.

## 2 Methods

Knowledge Discovery and Data Mining (KDD) methods require extensive data bases as an input. In cardiology we can acquire large volumes of heterogeneous patient data using different time scales. This is the basic reason that has led us to design of a data model, database architecture and processing steps that would allow efficient and transparent collection of heterogeneous data with different time scales in one storage and successive processing using different methods for different data types. It is process of original data set transform into strictly defined form. Basic diagram of integration, storage and KDD processes is shown in figure 1.

![Diagram of basic setting of the first processing step](http://meesoft.logicnet.de/)

**Fig 1. Diagram of basic setting of the first processing step**
A data model of the given problem is a very important part of the whole concept. The model has been derived from the following assumptions that were formulated based on consultations with medical doctors from the area of physiology and clinical cardiology. Basic unifying element of the record is the subject (patient). Basic entity having time specification is an event. It must be possible to attach to each event any number of signals (with respect to signal type, sampling rate, sensitivity, processing). A component of laboratory test results (type of tests, measured values and units) is formed analogously to results of each event. Information about long-term (e.g. height, diagnosis, personal and family anamneses, medication) and middle-term (e.g. weight, systolic and diastolic pressures, medication) state of subject is related to the measured subject and not to each event. All code values are defined in code lists – in addition to above mentioned values there can be included also values for signal segmentation, analysis and evaluation - characteristic points (generally marks of any event in signal), points for definition of complexes (the simplest complex is interval), analysis value (amplitude value, area value, etc). Schema of the model is shown in figure 2. The described model has been implemented as a relational database in Oracle database system 10.2.0 [1].

The software application has been divided into four basic parts. The first part includes forms for data entry on the measured subject. Values are entered by doctor’s or assistant’s selection of relevant code. This part contains computation and visualization of long-term and middle-term investigation trends in different time aggregation (minutes, hours and weeks). Example of two forms for investigation and description is shown in figure 7a) and 7b). Signal viewer opens in the second part. There is a defined protocol for specification of

![Fig 2. Data model (yellow color for patient part, cyan color for events, orange color for signals part, green color for point, complex and analysis part, grey color for laboratory, wheat color for evaluation part)](http://www.yworks.com)
leads (signal-reference terminal), thus forming unified lead list from various sources. In our case the most measured signal is multi-leads ECG (electrocardiograph). The main recording device is BioSemi ECG system [2]. However, data model considers more signal types. Therefore the application has the steps of signal loading and processing solved by means of the Interface (in accordance with Java definition) that allows use of various algorithm. Schematic circuit of protocol creation is shown in figure 3. The example of protocol instance definition is shown in figure 4.

![Fig 3. Diagram of protocol principle](http://meesoft.logicnet.dk/)

![Fig 4. Example of protocol creation in application](http://meesoft.logicnet.dk/)

We assume need of certain signal preprocessing. This step is also solved by means of interface in order to preserve defined variation of more signal sources. ECG signal is preprocessed according to recommendation [3]. Before we start signal preprocessing we need to distinguish between signal and single data values (usually data exported from a data source). Thus we define signal in our application as data that meets the following criteria: signal is any time-varying or spatial-varying quantity. It is expected that the information content in the signal is relevant to the whole event and/or it is possible to divide a signal into several events and assign more signals to one event (it is important to say that length of the signal is usually shorter than time interval between events). Within an event, there can be performed change of time resolution of the signal. It is possible to perform analysis that...
creates representative attribute(s) of signal information (descriptive feature/s). The goal of this definition is clear separation of data that have only a single value within one event or can be replaced by one value without any loss of information.

Signal viewer is basic part for visualization of measured signals with possibility of caliper application for measurement of intervals lengths and signals amplitudes, averaging and marking of characteristics points in the signal. The title characteristics points are defined as function points (e.g. ECG points), as mark of some event in signal (e.g. ablation) or beginning and end of some interval (e.g. fibrillation). The example of signals with marking of ECG characteristic points is shown in the figure 5. Data model supposes composition of complexes (interval in the simplest case) from defined points. The way of composition is given as combination of points with number of order in the complex and definition of the complex. This simple way provides general approach for creation of complexes based on function aspects and/or user definition. Complexes together with points represent basic elements for analysis. The analysis means extraction of information from signal into descriptive features.

To the part of signal displaying, there belong also functions of advanced visualization of the signal and its properties that can be also used for analysis of the signal. These additional components of signal viewer are mainly trends of statistic parameters, overview of calculated signal parameters (correlation, coherence, fractal dimension, etc.) and visualization of time and frequency signal properties (spectrum, spectrogram, continuous and discrete wavelet transform) [4].

In the third part mapping is computed and displayed. Grid of maps is defined in the protocol again. Mapping part includes four types of 2D maps (immediate, integral, isochronal, differential), overview of map signals sorted based on map grid and creation of greater number of maps for preview of surface potential distribution in time. The last mapping option is creation of frequency map in given number of bands. Example of immediate potential map is shown in figure 7d). The maps can be used for analysis in the same way as signals.

The last part of application is analysis of signal and its parts. The analysis is basic process for description of signal and feature extraction. The analysis has the same assumptions as signal readers and preprocessing steps. Thus results of analysis are transferred to model using the transforming interface that allows use larger number of various algorithms. It is important that the application is oriented mainly to integration of data (it does not need to include all analysis algorithms). Hence there is defined a possibility to transfer analysis results from
external source (with storage of type analysis and its parameters). The diagram of points, complex and analysis ordering is shown in figure 6. The results of analysis and information about long-time and middle-term values are used in evaluations, which are in many cases well-founded by rules from previous analysis or knowledge of experts. This process is not one-stage and we suppose that the evaluations table will contain several level of evaluation. Very simple example can be the following one: simple level of evaluation is “QRS complex is normal”, in rules it means the length of QRS > 0.06s and < 0.1 s; higher level of evaluation is “signal is normal”, in the rule from previous evaluations it means QRS complex is normal and QTc interval is normal, in basic rules it means length of QRS > 0.06s and < 0.1 s and length of QTc ≤ 0.40 s).

Fig 6. Diagram of preprocessing, points, complex and analysis ordering

Examples of individual application forms for entering information about basic patient information, diagnosis, medication and laboratory results, signal caliper and an example of potential map are shown in figure 7.

Fig 7. Example of application: a) Basic patient form, b) anamnesis, medication and results of laboratory tests form, c) signal caliper, d) immediate potential map
3 Results

The main result of this application is integration of basic patient data (age, gender, weight, height, systolic and diastolic pressures), information about diagnosis, anamnesis and medication, laboratory results and measured signals. Further we define process for signals segmentation in order to acquire time structure as interval and complex. From both these elements (generally points, complexes) we can perform analysis for extraction of signal description useful for evaluation of signal part and signal as complex respectively. Thus we acquire comprehensive overview of the measured subject and we can extract related information for more complex evaluation.

4 Conclusions

The aim of our project is to provide extended data processing methodology in the area of heart field observation and classification. We have developed a tool for long traces of synchronous heterogeneous data explanation and rapid evaluation of new data modalities. The field of cardiologic data is exactly the large area for that this approach is intended (especially in experimental electrophysiology). We are also well aware that this extended approach can be used in more fields, unlike general description of all areas, but as a frame for creation of such a description. Complex evaluations in neurology, neurophysiology or neuropsychology are examples of such fields.

Acknowledgments

This work was supported by the research program "Information Society" under grant No. 1ET201210527 "Knowledge-based support of diagnostics and prediction in cardiology" and currently is financed by the research program No. MSM 6840770012 “Transdisciplinary Research in the Field of Biomedical Engineering II” of the CTU in Prague, sponsored by the Ministry of Education, Youth and Sports of the Czech Republic.

References