Data Acquisition and Complex Systems Analysis

Kenneth A. Loparo

Department of Electrical Engineering and Computer Science

Case Western Reserve University

September 26, 2009
Current ICU Monitoring

Monitors/Sensors

Patient

Trend Visualization and Display

Process
Proposed ICU Monitoring

Monitors/Sensors

Plug/Play Std Interface

Data Warehouse

Data Aggregation and Feature Extraction

Visualization

Process
Phase 1: Data Integration

Development of proof of concept data collection system, built from “Commercial Off The Shelf” components, that will connect with current patient monitoring devices.

The goal will be to collect and integrate not only the parametric data but also acquire the underlying waveforms at high sampling rates.
Phase 2 will be divided into two phases:

**Phase 2a:** Develop trend detection, feature extraction and data processing software.

**Phase 2b:** Develop a user-friendly and intuitive interface for signal quantification, interpretation and display.
Fundamentals of Data Acquisition

Analog Signal → Sampling → Sampled Signal → Digital Signal

Fourier Transform

Sampling

F

-F_s

F_s = 1/T_s
Fundamentals of Data Acquisition

Analog Signal → Sampling → Sampled Signal → Digital Signal

Fourier Transform

Sampling

Fourier Transform

Aliasing due to undersampling

\( F_s = \frac{1}{T_s} \)
Complex Systems Analysis

Complexity Analysis: To distinguish and quantify complexity from different sources, e.g. stochastic and deterministic

Variability Analysis: Quantify how patterns are changing over time

Applications:

• Engineering
• Physical and social sciences
• Biology
• Medicine
  – Physiological time series data
    EEG, respiration, heart rate, blood pressure, sleep/awake cycles, etc.
Classical Example: Heart Rate Variability

• Heart rate variability (HRV) refers to the beat-to-beat alterations in heart rate.

• Under resting conditions, the ECG of healthy individuals exhibits variability in the R-R intervals.

• The natural rhythmic phenomenon (respiratory sinus arrhythmia-RSA) fluctuates with the phase of respiration: cardio-acceleration during inspiration and cardio-deceleration during expiration.
  • Leading to quantifiable changes in temporal R-R interval patterns as well as the dynamic coupling between cardio-respiratory system.

• The objective is to quantify how these patterns are affected by health and disease, and use this information for disease diagnosis and prognosis and to improve outcome through therapeutic intervention.
Interval Tachogram: HRV signal defined as the series of R-R interval durations as a function of the interval number. If $t_n$ is the occurrence time of the n-th R-wave, the interval tachogram is $IT(n)$, and is given by:

$$IT(n) = (t_n - t_{n-1})$$

Instantaneous Heart Rate: HRV signal defined as the series of the reciprocal of R-R interval durations as a function of time. If $t_n$ is the occurrence time of the n-th R-wave, the instantaneous heart rate $IHR(t)$ is given by:

$$IHR(t) = \sum_{t=-\infty}^{\infty} \frac{1}{(t_n - t_{n-1})} \delta(t - t_n)$$
Heart Rate Dynamics
Heart Rate Dynamics

Congestive heart failure

Normal Rhythm: appears nonstationary and nonlinear

Atrial Fibrillation: appears random and uncorrelated
Why Study Complex Systems?

- Physiological and biological systems in health and disease exhibit an extraordinary range of temporal behaviors and structural patterns that may defy understanding based on concepts and principles of linear systems and analysis.
- Concepts and computational tools derived from the contemporary study of complex systems including nonlinear dynamics, fractals and variability are having an increasing impact on biology and medicine.
- Such concepts and computational tools can provide more information leading to a deeper understanding of a system and the underlying sources that are responsible for the irregular behavior.
Linear verses Nonlinear Systems

• Two key principles of linear systems are proportionality and additivity
• The overall behavior of a linear system is described by a linear combination of its constituent parts
• The proportionality and/or additivity principles fail to capture the behavior of the simplest nonlinear system such as the logistic map: \( y = ax(1-x) \)
• Nonlinearity is necessary for complex dynamics
• Complex (nonlinear) systems are composed of multiple subunits that cannot be understood by analyzing these components individually
• Complex (nonlinear) systems need to be studied as an interconnected system
A dynamical consists of 2 parts:
- State
- Dynamics

The state is defined as the information necessary at a given time instant to define further outputs of the system.

The dynamics are a set of rules that define how the state evolves over time.

A dynamical system evolves on a state space (also called phase space).

The state is represented by a vector in the state space.

A mathematical description of a dynamical system is
- System of difference equations for the discrete case
- Systems of differential equations for the continuous case

The motion of the system is called a trajectory or orbit of the dynamical system.
Complex Dynamics: Discrete-Time Systems

- Dynamical system: $x_{k+1} = f(x_k)$
- Initial conditions: $x_0$
- Orbit: $\{x_k\}_{k \geq 0}$ where $x_k = f^k(x_0)$

- Complex dynamics and chaos:
  - Sensitive dependence on IC’s (unpredictability)
  - Topological transitivity (indecomposability-dense orbits)
  - Dense set of periodic points (element of regularity)
Nonlinear, chaotic systems can produce very irregular data; similar but distinct from a stochastic system.

The rapid loss of predictability of chaotic systems is due to a phenomenon called *sensitive dependence on initial conditions*.

For a certain class of dynamical systems, once transients are over, the trajectory of the system approaches a subset of the state space called an *attractor*.

Attractors of dissipative chaotic systems generally are strange attractors—complicated geometrical objects that exhibit fractal structure.

Examples of deterministic chaotic systems: Logistic Map, Lorenz attractor
Example: Logistic Map-Iterates

\[ x_{k+1} = rx_k (1 - x_k) \]
Figure 3.3 Logistic map with $r = \frac{7}{10}$: (a) The time series quickly stabilizes to a fixed point. (b) The state space of the same system shows how subsequent steps of the system get pulled into the fixed point.
Figure 3.4 Logistic map with $r = \frac{3}{4}$: (a) The time series quickly stabilizes to a period-2 limit cycle. (b) The state space of the same system shows how subsequent steps of the system get pulled into the limit cycle. (c) The state space of the same system but with only the converged values for $x_i$ plotted, so as to clearly show the limit cycle’s location.
Example: Logistic Map-Iterates

![Logistic Map-Iterates](image)

**Figure 3.5** Logistic map with $r = \frac{35}{44}$. (a) The time series quickly stabilizes to a period-4 limit cycle. (b) The state space of the same system. (c) The state space of the same system but with only the converged values for $x_n$ plotted.
Example: Logistic Map-Iterates

Figure 3.6  Logistic map with r = 1: (a) The time series is chaotic and has the appearance of noise. (b) The state space of the same system, which illustrates how the system’s trajectory visits every local region. (c) The state space of the same system with only four steps plotted, so as to show how small differences turn into larger differences.

Figure from The Computational Beauty of Nature: Computer Explorations of Fractals, Chaos, Complex Systems, and Adaptation. Copyright © 1998-2000 by Gary William Flake. All rights reserved. Permission granted for educational, scholarly, and personal use provided that this notice remains intact and unaltered. No part of this work may be reproduced for commercial purposes without prior written permission from the MIT Press.
Example: Logistic Map-Routes to Complex Dynamical Behavior
For a smooth (i.e., nonfractal) line, an approximate length \( L(r) = N r \);
\( N \) = min number of segments of length \( r \) needed to cover the line.

The area \( A \) or the volume \( V \) of a nonfractal object is the limit of an integer power law of \( r \):

\[
A = \lim_{r \to 0} N r^2 \quad V = \lim_{r \to 0} N r^3
\]

where the exponent = the Euclidean dimension of the object.
For fractal objects: as \( r \) tends to 0, finer and finer details of the fractal appear and the product \( N r^E \) may diverge to infinity.

A real number \( D \) exists so that the limit of \( N r^D \) stays finite. This exponent is called the Hausdorff Dimension:

\[
D_H \equiv \lim_{r \to 0} \frac{\log N}{\log(1/r)}
\]
The **Correlation Dimension** $D_2$ of a time series $\{x_k\}$ measures the active degrees of freedom of the system on an attractor:

$$D_2 \equiv \lim_{r \to 0} \frac{\log C(r)}{\log(r)}$$

$$C(r) = \lim_{N \to \infty} \frac{1}{N^2} \sum_{i=1, j=1}^{N} \Theta(r - |x_i - x_j|)$$

Where $\Theta$ is the Heavyside function and for small $r$,

$$C(r) \approx r^{D_2}$$
Example: $D_2$ of Lorenz Attractor

\[
\frac{dx(t)}{dt} = \sigma(x - y) \\
\frac{dy(t)}{dt} = x(\rho - z) - y \\
\frac{dz(t)}{dt} = xy - \beta z
\]
Tests for Nonlinearity and Complex Dynamics: Surrogate Data Testing

- The method of surrogate data is one of the most popular tests that are used to examine the evidence of nonlinearity in a time series.
- The null hypothesis of the surrogate data test is to statistically determine whether or not a time series is generated from a Gaussian linear stochastic process using computational measures (e.g., correlation dimension) of both the original time series and its surrogate data.
- Surrogate data can be simply generated by randomly shuffling the original time series—can lead to erroneous results.
- Another surrogate data generation method is called the method of iteratively refined surrogates.
- In this method, surrogates are iteratively refined to correct deviations in the power spectrum and distribution from the original time series.
- Surrogate data generated by the method of iteratively refined surrogates have approximately the same power spectrum and amplitude distribution as the original time series.
Example: $D_2$ of Surrogate Data
Let $X$ and $Y$ be dynamical systems that generate scalar time series $x$ and $y$.

A geometrical object in the state space reconstructed from both dynamical systems $X$ and $Y$ compared to their individual geometrical objects provides information about the interconnection of the dynamical systems $X$ and $Y$.

There are 3 steps for measuring the synchronization between the dynamical systems $X$ and $Y$:

- **Step 1:** Compute the complexity of the dynamical systems $X$ from the time series $x$, and the complexity of the dynamical systems $Y$ from the time series $y$.
- **Step 2:** Compute the complexity of the mutual geometrical object of both dynamical systems reconstructed by concatenating the state vectors of $X$ and $Y$.
- **Step 3:** Compute the coupling index of the dynamical systems $X$ and $Y$ is defined by

$$\Gamma_{X,Y} = \frac{D_2(X) + D_2(Y)}{D_2(X,Y)}$$

The coupling index, called *dynamic complexity coherence*, specifies the strength of the coupling between the dynamical systems $X$ and $Y$. 
Example: Synchronized System

\[ D(X) = 1 \]
\[ D(Y) = 1 \]
\[ D(X,Y) = 1 \rightarrow \Gamma_{X,Y} = 2 \]
Example: Unsynchronized System

\[ D(X) = 1 \]

\[ D(Y) = 1 \]

\[ D(X,Y) > 1 \rightarrow \Gamma_{X,Y} < 2 \]
Example: Synchronized Lorenz System
Fractals and Self-Similarity

\[ y(t) = a^\alpha y\left(\frac{t}{\alpha}\right) \]

\( \alpha \) is the self-similarity parameter
Detrended Fluctuation Analysis

Method for quantifying the correlation properties in nonstationary time series based on the computation of a scaling exponent $d$ by means of a modified root mean square analysis of a random walk.

$$y(k) = \sum_{i=1}^{k} [x(i) - M]$$

$y(k)$ is divided into segments of equal length $n$, for various scales $n$, where for each $n$ $y_n(k)$ is computed from a linear least-squares fit to the data in each box.
Detrended Fluctuation Analysis

The integrated time series is detrended by subtracting the local trend \( y_n(k) \), and the root-mean square fluctuation of the detrended series, \( F(n) \), is computed using the formula:

\[
F(n) = \sqrt{\frac{1}{N} \sum_{k=1}^{N} \left[ y(k) - y_n(k) \right]^2}
\]

This computation is repeated over all time scales (box sizes) to provide a relationship between \( F(n) \) and the box size \( n \). Typically, \( F(n) \) will increase with box size \( n \). A linear relationship on a log-log graph indicates the presence of scaling (self-similarity)—the fluctuations in small boxes are related to the fluctuations in larger boxes in a power-law fashion. The slope of the line relating \( \log F(n) \) to \( \log n \) determines the scaling exponent (the self-similarity parameter) \( d \) where:

- \( d \) is related to the \( 1/f \) spectral slope: \( \alpha = 2d - 1 \) and
- If \( d = 0.5 \), the time-series is uncorrelated (white noise).
- If \( d = 1.0 \), the correlation of the time-series is the same as \( 1/f \) noise;
- If \( d = 1.5 \), time series behaves like Brownian motion (random walk).
Detrended Fluctuation Analysis

The time series is integrated

\[ y(k) = \sum_{i=1}^{k} \left[ B(i) - B_{ave} \right] \]

The integrated data is divided into equal segments of size ‘\( n \)’. The trend in each segment is removed using a linear least squares fit. The Root Mean Square (RMS) fluctuation \((F)\) is calculated. This process is repeated to obtain \(F\) as a function of \(n\).

\[ F(n) = \sqrt{\frac{1}{N} \sum_{k=1}^{N} [y(k) - y_n(k)]^2} \]

The gradient of the bi-logarithmic plot of \(F\) vs. \(n\) is estimated.
DFA applied to HRV

Scaling Analyses of the Healthy Heartbeat

(a) DFA Analysis
- Healthy subject, $\alpha=1.04$
- Randomized control, $\alpha=0.51$

(b) Spectral Analysis
- Healthy subject, $\beta=1.03$
- Randomized control, $\beta=0$

Original time series data is randomized
DFA: Aging and Disease

Altered Fractal Scaling with Heart Failure and Aging

- **Healthy Young**
- **Healthy Elderly**
- **Heart Failure**

*white noise, slope=0.5*
*1/f noise, slope=1*
*random walk, slope=1.5*
Variability Analysis Techniques

Time domain analysis
- Mean, variance, etc

Frequency domain analysis
- Windowed spectral analysis, PSD, Periodogram, Wavelets, etc

Entropy measures:
- Shannon entropy, approximate entropy, sample entropy, interval entropy, and spectral entropy

Linear stochastic signal models
- Regression, correlation analysis, coefficient of variation, time domain modeling (AR, ARMA, etc)

Scale independent or power-law techniques:
- Correlation dimension, Detrended Fluctuation Analysis, …
Define: $v(n) = [x(n), x(n+1), \ldots x(n+m-1)]^T$ from the signal samples $x(n)$. $D(i,j) = \text{the maximum difference in the scalar components of } v(i) \text{ and } v(j)$, and $N_{m,r}(i) = \text{the number of vectors } j$ (with $j \leq N-m+1$) such that the distance between the vectors $v(j)$ and the generic vector $v(i)$ (with $i \leq N-m+1$) is less than $r$, i.e $D(i,j) \leq r$. Here $r$ is a fixed parameter that determines the "tolerance" of the comparison. Define $C_{m,r}(i) = \text{the probability of finding a vector that differs from } v(i) \text{ with distance less than } r$: $C_{m,r}(i) = \frac{N_{m,r}(i)}{N-m+1}$ Let: $F_{m,r} = \sum_{i=1}^{N-m+1} \ln(C_{m,r}(i))$ Then: $ApEn_{m,r} = F_{m,r} - F_{m+1,r}$
ApEn changed to 1.69, 1.50 and 1.51 respectively after randomization of the samples.
Automatic Tracking of Scale-Dependent Signal Characteristics using Change Point Detection

\[ \dot{x} = \frac{dx(t)}{dt} \]

is approximated using three-point Lagrangian approximation of the derivative:

\[ \dot{x}(t_k) = \frac{x(t_{k+1}) - x(t_{k-1})}{2h} \]

Change point detection [Brodsky et al. 1999] identifies the most statistically invariant segments

\[ h = t_{k+1} - t_k = t_k - t_{k-1} \]
1) Ti-Hr= Interval from onset of Inspiration to the next EKG R-wave.

2) Te-Hr= Interval from onset of Expiration to the next R-wave.

3) Hr-Ti= Interval from the R-wave to the onset of Inspiration.

4) HR-Te= Interval from the R-wave in Inspiration to Expiration.

EKG=RED
FLOW=Blue
The concept of Shannon Entropy, a central concept in information theory, is a measure of uncertainty that has also been used as a measure of information content. The entropy of a discrete random variable with probability $p_i$ is given by:

$$- \sum_{i=1}^{N} p_i \log_2 p_i$$

where $N$ is total number of events. For a continuous random variable $x$, we have:

$$H(x) = \int_{S} f(x) \log_2 f(x) dx$$

The entropy of an interval is defined to be the continuous Shannon entropy of the interval of interest.
Event Coupling Detection and Quantification Algorithm

Detection of R-waves and the onset of inspiration and expiration.

Calculate the desired time intervals between heart activation and inspiration or expiration.

Generate 19 surrogate heart rate signals to meet the statistical significance as follows:

- Detection of R-R intervals.
- Generation of surrogate R-R. The surrogate R-R time series have the same mean, standard deviation and histogram of the original R-R time series. Also, the autocorrelation of the surrogate R-R is almost the same as original R-R.
- Surrogate heart rate generation by cumulative sum of surrogate R-R intervals.
- Entropy analysis of the intervals computed from both the original and surrogate data of heart activation, inspiration and expiration.

Rank ordering and counting the number of entropy values for the intervals of the surrogate data that exceed those of the original data, to quantify, in a statistical sense, the existence of non-linear coupling in the data.
# Real-Time Feature Extraction Results

## Total Recording Time (Sec)

<table>
<thead>
<tr>
<th>Subject</th>
<th>Total Recording Time (Sec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subject 1</td>
<td>1425</td>
</tr>
<tr>
<td>Subject 2</td>
<td>1208</td>
</tr>
<tr>
<td>Subject 3</td>
<td>2949</td>
</tr>
</tbody>
</table>

## Detected R-Wave

<table>
<thead>
<tr>
<th></th>
<th>Subject 1</th>
<th>Subject 2</th>
<th>Subject 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Detected R-Wave</td>
<td>2129</td>
<td>1659</td>
<td>2657</td>
</tr>
</tbody>
</table>

## Average Heart Rate (B/Min)

<table>
<thead>
<tr>
<th>Subject</th>
<th>Average Heart Rate (B/Min)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subject 1</td>
<td>89.64</td>
</tr>
<tr>
<td>Subject 2</td>
<td>82.40</td>
</tr>
<tr>
<td>Subject 3</td>
<td>54.06</td>
</tr>
</tbody>
</table>

## Average Breaths Per Minute

<table>
<thead>
<tr>
<th>Subject</th>
<th>Average Breaths Per Minute</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subject 1</td>
<td>9.17</td>
</tr>
<tr>
<td>Subject 2</td>
<td>5.31</td>
</tr>
<tr>
<td>Subject 3</td>
<td>6.43</td>
</tr>
</tbody>
</table>

## DFA, SD1C, SD2C, SD1C to SD2C Ratio, Conditional Entropy

<table>
<thead>
<tr>
<th></th>
<th>DFA</th>
<th>SD1C</th>
<th>SD2C</th>
<th>SD1C to SD2C Ratio</th>
<th>Conditional Entropy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subject 1</td>
<td>0.74</td>
<td>0.0346</td>
<td>0.1485</td>
<td>0.23</td>
<td>1.21</td>
</tr>
<tr>
<td>Subject 2</td>
<td>0.78</td>
<td>0.0091</td>
<td>0.0645</td>
<td>0.14</td>
<td>1.17</td>
</tr>
<tr>
<td>Subject 3</td>
<td>0.61</td>
<td>0.0586</td>
<td>0.1430</td>
<td>0.40</td>
<td>1.30</td>
</tr>
</tbody>
</table>

## Interval Entropy

<table>
<thead>
<tr>
<th></th>
<th>R-R</th>
<th>Total Ti</th>
<th>Total Te</th>
<th>Ti</th>
<th>Te</th>
<th>HrTi</th>
<th>HrTe</th>
<th>TiHr</th>
<th>TeHr</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subject 1</td>
<td>-1.31</td>
<td>3.47</td>
<td>3.44</td>
<td>2.36</td>
<td>3.84</td>
<td>-0.57</td>
<td>-0.80</td>
<td>-0.59</td>
<td>-0.75</td>
</tr>
<tr>
<td>Subject 2</td>
<td>-2.45</td>
<td>3.95</td>
<td>3.94</td>
<td>3.23</td>
<td>4.48</td>
<td>-0.60</td>
<td>-0.58</td>
<td>-0.66</td>
<td>-0.51</td>
</tr>
<tr>
<td>Subject 3</td>
<td>-1.18</td>
<td>3.38</td>
<td>3.36</td>
<td>2.35</td>
<td>3.78</td>
<td>0.15</td>
<td>0.22</td>
<td>0.12</td>
<td>0.27</td>
</tr>
</tbody>
</table>
## Alpha Delta Monitoring Criteria

- Very simple and easy to use
- Alpha 8-12 Hz
- Delta 1.5-2.5 Hz
- FFT calculation based on 30 Sec epoch
- Display as 15 minute ratios
- Normalization
- Real-time processing
Channel Selection and Display

Selecting desired Channels

- Symmetric channels (Left & Right Hemisphere).

Display the difference in the power ratio between the left and right hemispheres
Real-Time Alpha-Delta Monitoring