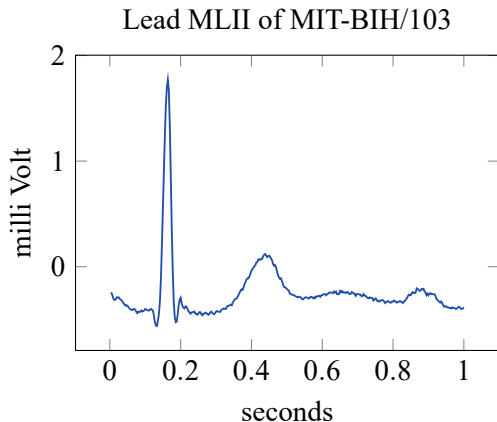# What is an ECG?



Lead MLII of MIT-BIH/103

Figure 1: ECG of one heartbeat

- electrocardiogram (ECG or EKG) records the heart's electrical activity

- contains up to 12 simultaneous measurements – the leads [moody1992a]
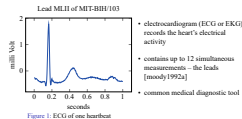
- common medical diagnostic tool

MSAX for ECG Analysis
└─Introduction
  └─ECG Basics
    └─What is an ECG?

2021-05-02

Figure 1: ECG of one heartbeat

- muscle contractions caused by electric pulses
- electric pulse can be measured on the skin
- the measuring things are called electrodes
- electrodes form leads (need 2 to measure anything)
- they have specific positions and names
- 12 leads is the modern standard
- most types of heart disease can be detected
- diagnosis and analysis is performed by trained cardiologists
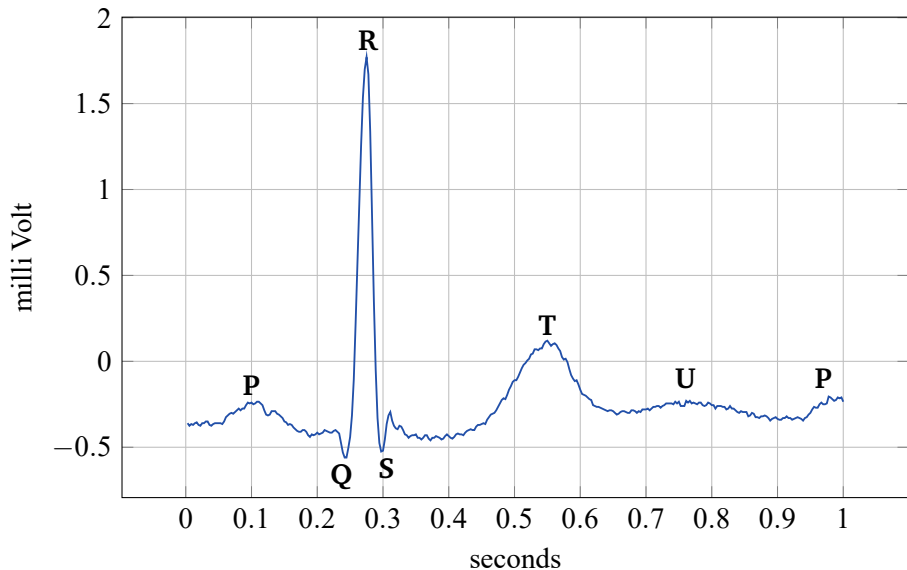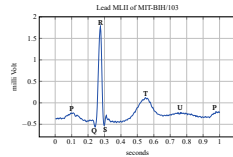- datasets available online; contain 2 or more leads (the most significant ones)

Figure 2: Annotated ECG of one heartbeat

Figure 2: Annotated ECG of one heartbeat

- P wave: atria depolarizing / blood entering the heart
- QRS complex: ventricular depolarization / heart contraction pumping blood
- T: return of ventricle to polarized state
- U: present in 25%, maybe some feedback
- P wave: atria depolarizing / blood entering the heart
- ST-segment: significant, depression, elevation, slope show ischaemia
- R-R interval: shows rhythm and thus arrhytmia etc

# ECGs as Time Series

### Definition

A discrete multivariate time series is a sequence at discrete points in time that has $n$ values at each of these points. If $n = 1$, the series is univariate and if $n > 1$, it is multivariate.

- digital ECGs are discrete multivariate time series:
    - have $> 1$ value at each point

    - recorded at discrete, evenly spaced time points
  [anacleto2020]

- time series analysis methods can be applied to ECGs

### ECGs as Time Series

**Definition**

A discrete multivariate time series is a sequence at discrete points in time that has $n$ values at each of these points. If $n = 1$, the series is univariate and if $n > 1$, it is multivariate.

- digital ECGs are discrete multivariate time series:
  - have $> 1$ value at each point
  - recorded at discrete, evenly spaced time points [amacleto2020]
- time series analysis methods can be applied to ECGs

- modern ECGs have at least 2, most have 12
- digital ones have set sampling frequencies, even the machines have set frequencies
- multivariate: measure more than 1 lead per time point
- discrete: set sample frequency in the machines
- discrete: because measured at discrete moments in time
- time series: they are data measured at equal time intervals
- $n$ measurements per point in time (i.e. leads)
- $n = 1$ is univariate, $n > 1$ is multivariate

# ECG Analysis

- standard method: manual analysis by cardiologist

- recently: automated (computer-assisted) ECG analysis

- multiple stages:(1) signal acquisition, filtering; (2) data transformation, processing; (3) waveform recognition, feature extraction; (4) classification [kligfieldpaul2007]

- common methods: FFT, DWT, ANN, kNN,…

- relatively new methods are SAX, MSAX, and HOTSAX

## ECG Analysis

- standard method: manual analysis by cardiologist
- recently: automated (computer-assisted) ECG analysis
- multiple stages: (1) signal acquisition, filtering; (2) data transformation, processing; (3) waveform recognition, feature extraction; (4) classification [kligfieldpaul2007]
- common methods: FFT, DWT, ANN, kNN,…
- relatively new methods are SAX, MSAX, and HOTSAX

- is relatively slow; time is of the essence
- lots of training required
- error prone
- maybe not feasible for long ECGs
- can speed up process
- can pick up details humans miss
- digitizing paper ECGs or recording digital ones
- filtering to remove various types of noise
- reduce complexity of the data
- balance between accuracy and complexity needed

# SAX, MSAX, and HOTSAX

- Symbolic Aggregate Approximation (SAX) creates a simplified, symbolic representation [lin2003]    [zhang2019]

- Multivariate SAX (MSAX)[anacleto2020] expands SAX to multivariate time series

- HOTSAX is a discord discovery algorithm that has been used with SAX

→ using HOTSAX with MSAX on ECGs should increase the accuracy of discord detection compared to HOTSAX with SAX

2021-05-02

MSAX for ECG Analysis
└─Introduction
    └─ECG Analysis
       └─SAX, MSAX, and HOTSAX

SAX, MSAX, and HOTSAX

- Symbolic Aggregate Approximation (SAX) creates a simplified, symbolic representation [lin2003]   [zhang2019]
- Multivariate SAX (MSAX)[anacleto2020] expands SAX to multivariate time series
- HOTSAX is a discord discovery algorithm that has been used with SAX
→ using HOTSAX with MSAX on ECGs should increase the accuracy of discord detection compared to HOTSAX with SAX

- represent ecg as letters that still mean the same thing
- is guaranteed to behave like the original data
- works on univariate time series, has been used on ECGs
- takes the correlation between ecg leads into account
- mention that MSAX to ECGs in particular is new
- mention that HOTSAX with MSAX is new