EEG classification and cross-validation using the BCILAB toolbox: intro

Christian A. Kothe
SCCN/UCSD

Toolbox download link:

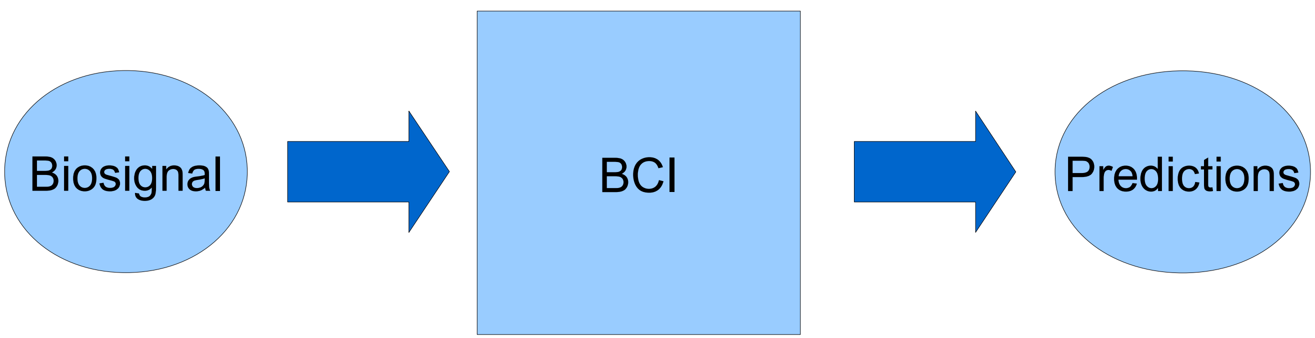
ftp://sccn.ucsd.edu/pub/bcilab/

Outline

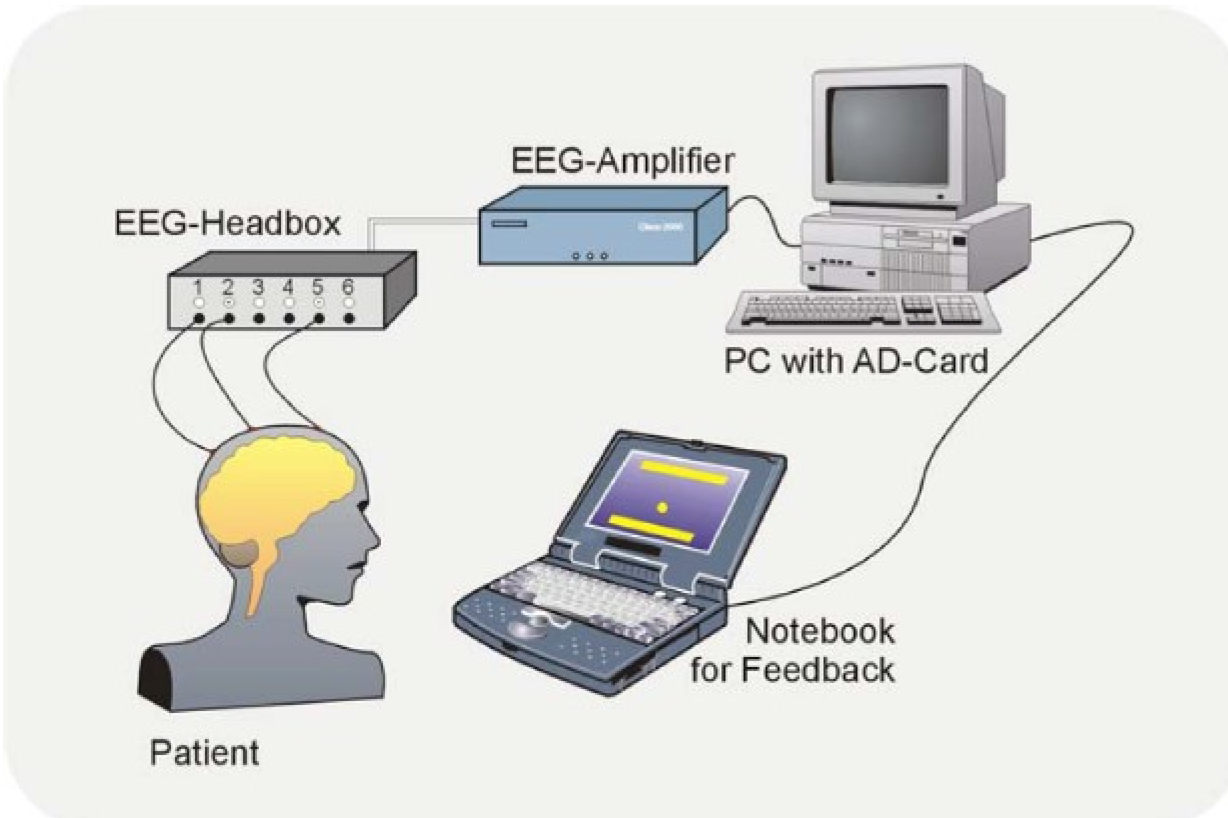
- Definition
- Motivation
- Operational basis
- Model structure
- Model learning
- Toolbox overview

What is a Brain-Computer Interface?

All-inclusive definition: A BCI is a system which takes a biosignal measured from a person, and predicts (in real time) some aspect of the person's cognitive state.



What is a Brain-Computer Interface?



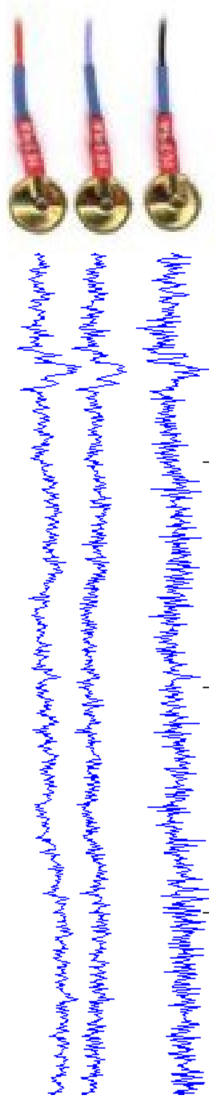
Schmidt, 2003

Areas of Research

- **Clinical**: Communication and control devices for the severely disabled (e.g., spelling devices in late-stage ALS)
- **HCI**: User-state monitoring for intelligent assistive systems, complex/demanding operational environments, etc.
- **Entertainment**: Computer game controllers (e.g., Emotiv, NeuroSky)
- **Neuroscience**: Pattern recognition and machine learning on brain signals, computational expert knowledge?



Operational Basis

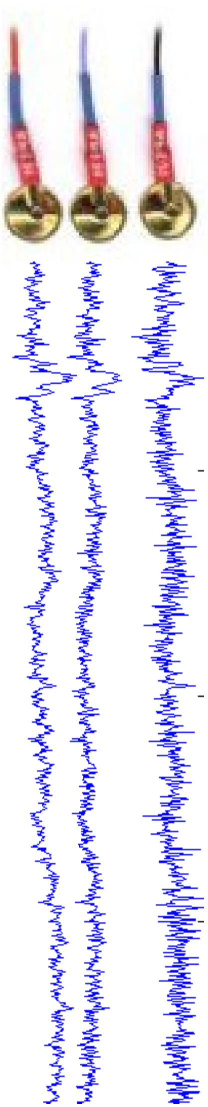


← Sensors

Time

← Sensor State

Operational Basis

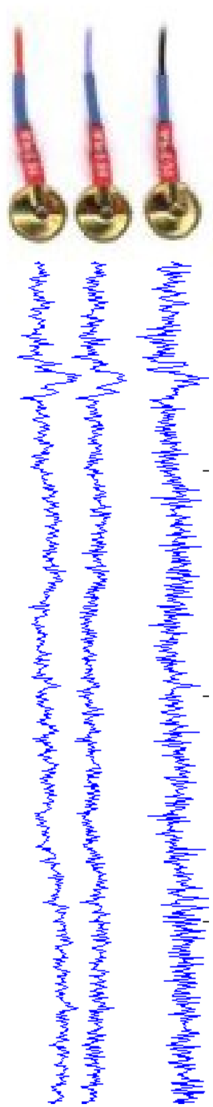


Brain →

Time

Brain State →
(too complex to show)

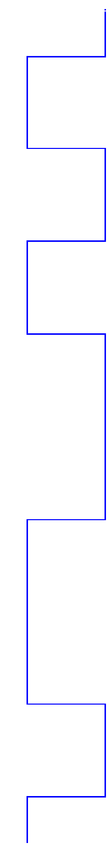
Operational Basis



?

Time

Abstract user state of interest →
(e.g., surprised/not surprised)

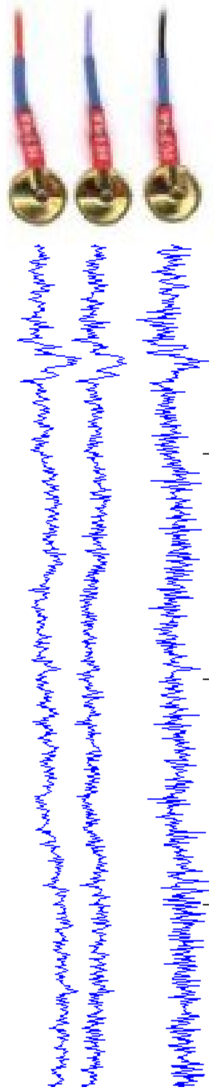


Operational Basis

- We postulate the existence of a certain aspect of the user's state (e.g., likelihood of being surprised, #items in working memory, type of imagined movement)
- To be able to estimate it from sensor readings, we require that a relationship between the two exists
- Physics provides some level of justification

Operational Basis

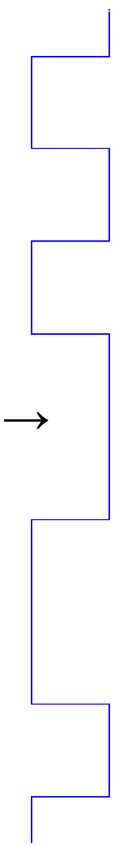
Time



← Physically →
related

← Physically →
related

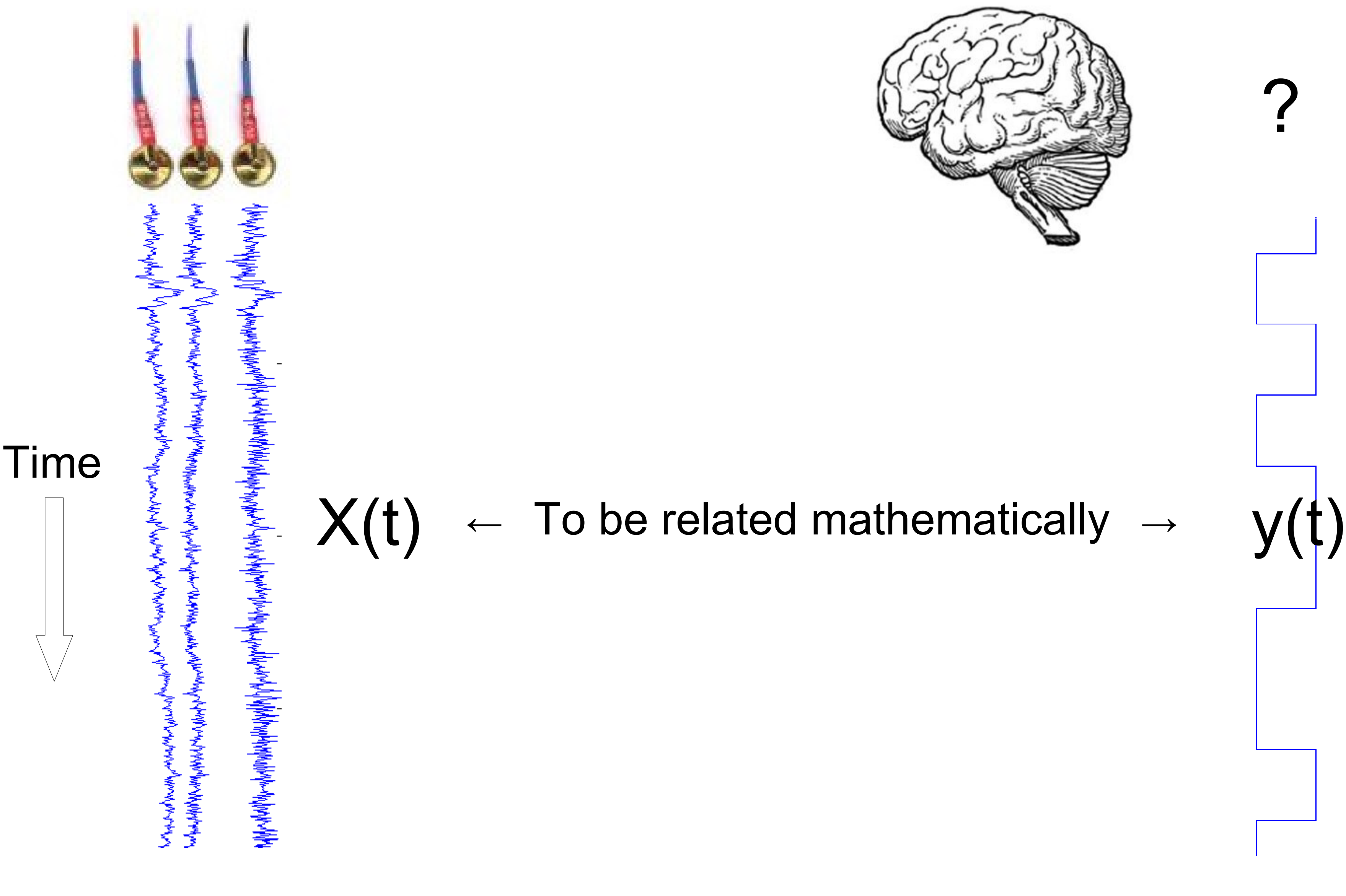
?



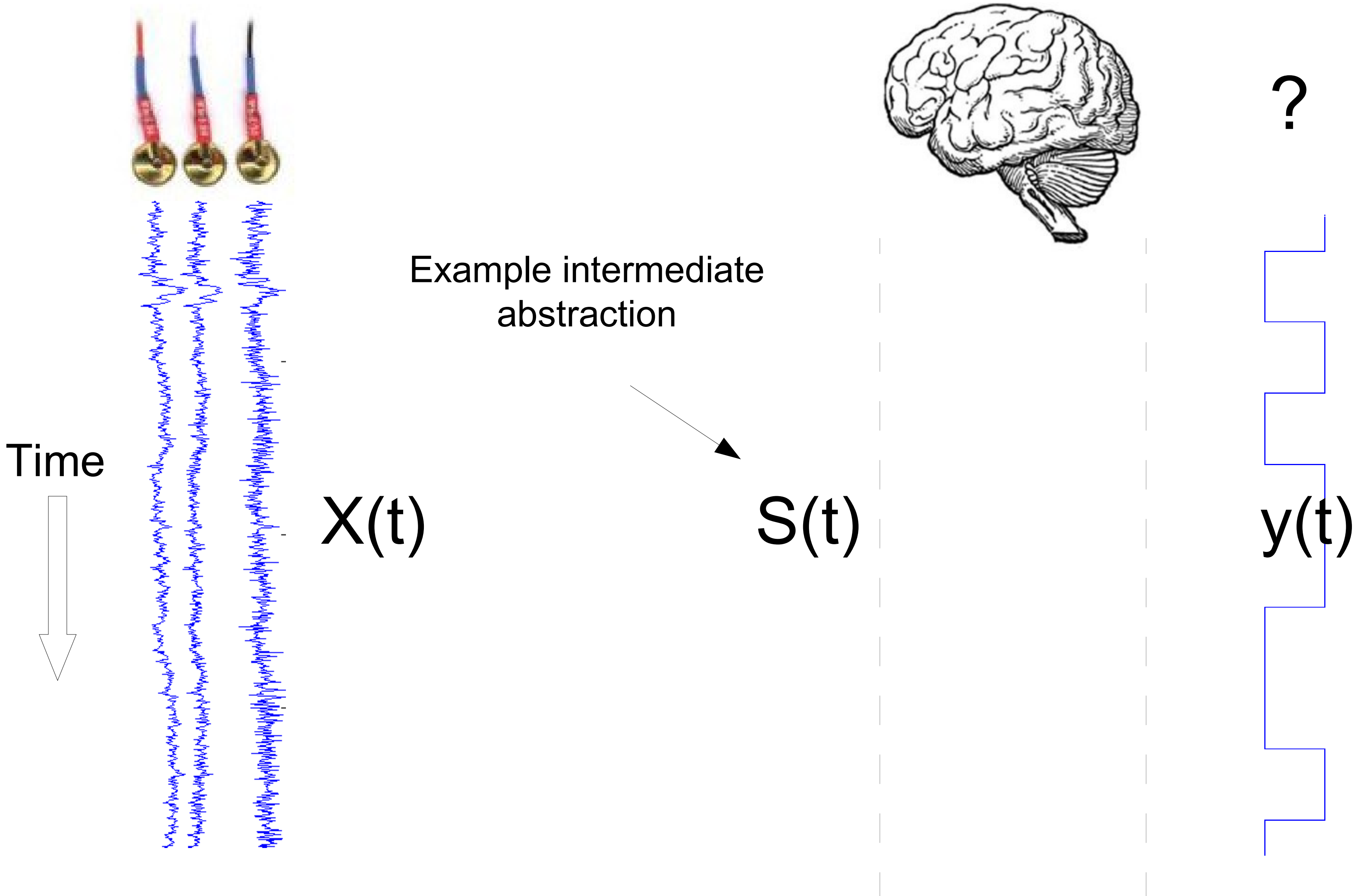
Operational Basis

- The task is to find a mathematical mapping between sensor signals and user state of interest that is accurate, robust, and computationally tractable
- Usually, many such mappings could be defined for any given question about cognitive state, depending on assumptions made
- Example: Spectral properties for oscillatory processes, time averages for slow-wave phenomena, partial directed coherence for information flow aspects

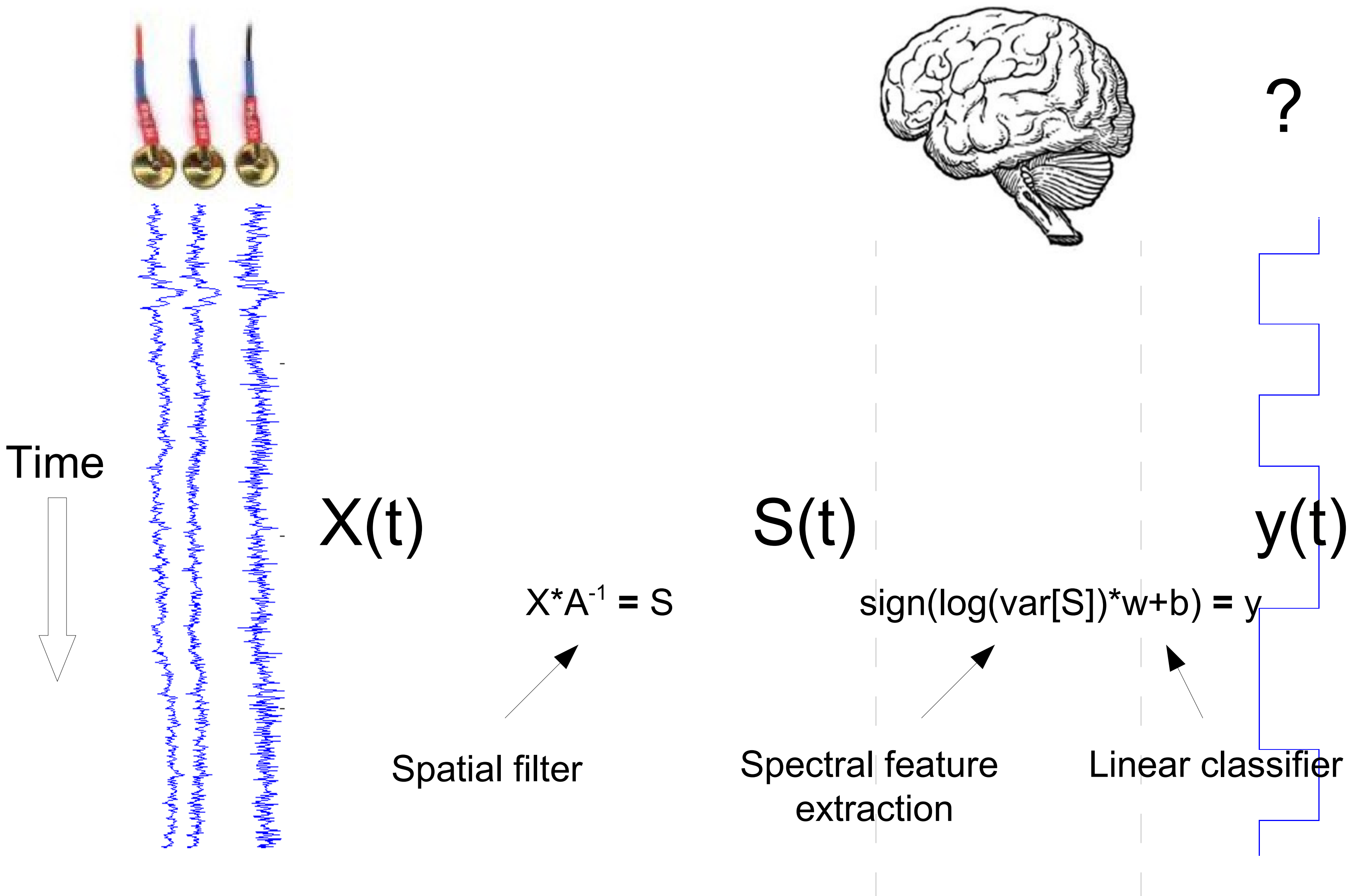
Defining a BCI approach



Defining a BCI approach



Defining a BCI approach

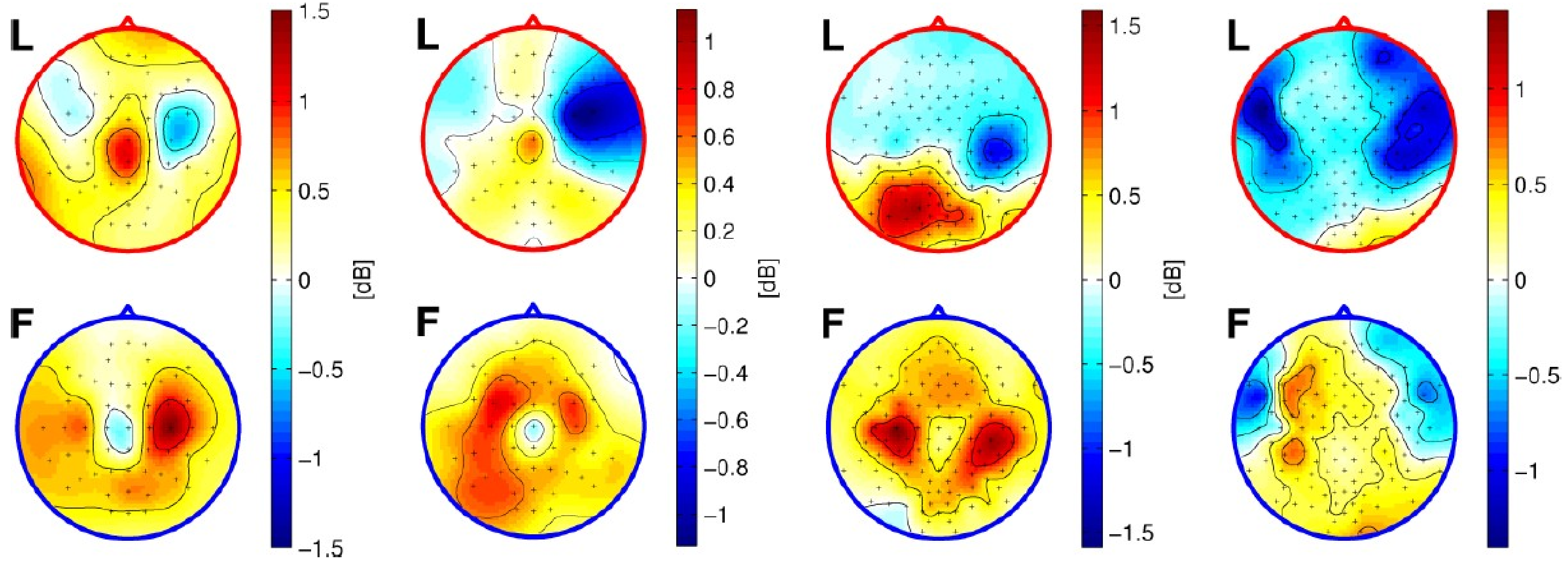


Defining a BCI approach, ct'd

- Problem: The mapping involves parameters (here: A^{-1}, w, b) that need to be obtained
- Optimal values for many such parameters vary drastically over people (and/or sessions, tasks, etc.)
- Thus, in many cases they need to be learned on the basis of source data
- Sometimes easy, but frequently hard

The problem of variability

- Spatial filters across 4 persons for predicting imagined movements (Left hand vs. Foot):



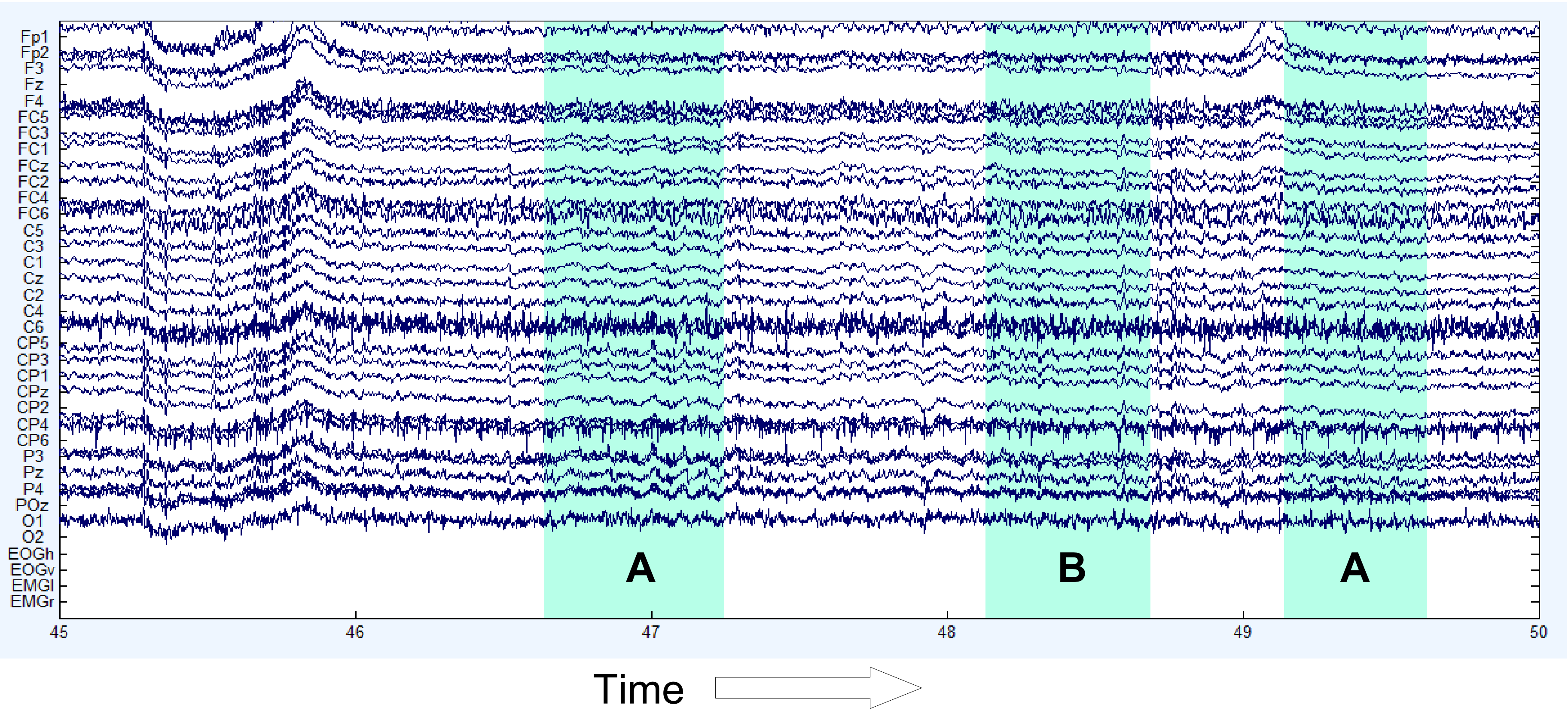
Blankertz et al., 2007

The problem of variability

- Many data sources can in principle help (e.g., MR images)
- Best known type of data source: *Example data*
  - Dedicated “training” / “calibration” session
  - Recording of both bio-signals X and target variables y across multiple trials
  - y known by construction of the experimental paradigm (e.g., via instructions, stimuli, ...)

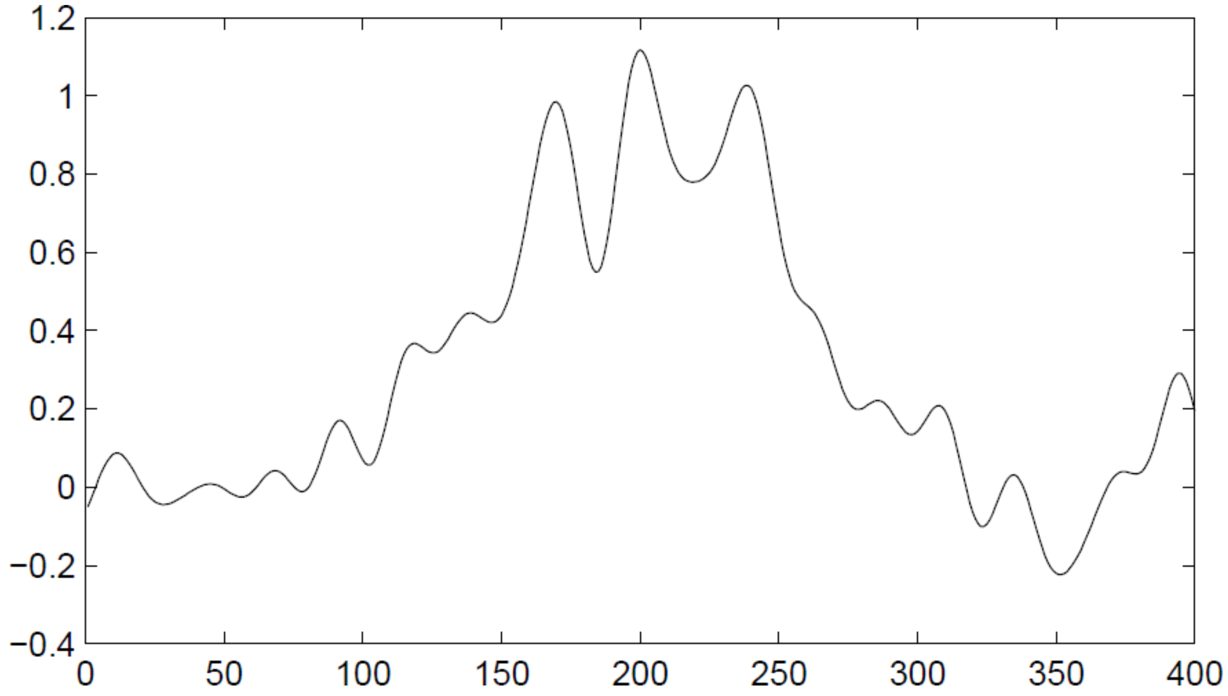
Concrete Example

- Typical calibration data contains segments (trials) for which there is information on the cognitive state of interest (e.g., via stimuli)

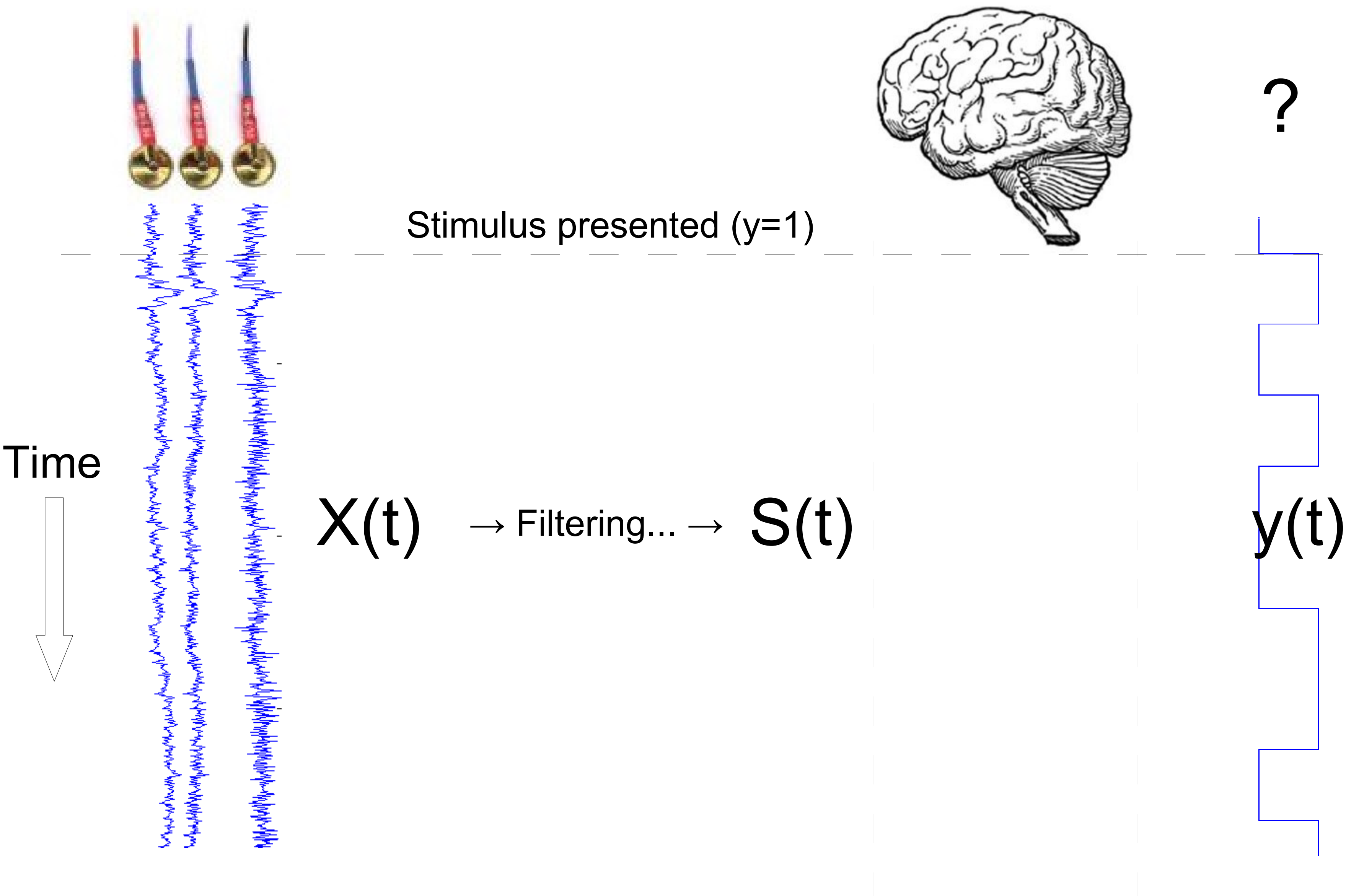


Concrete Example

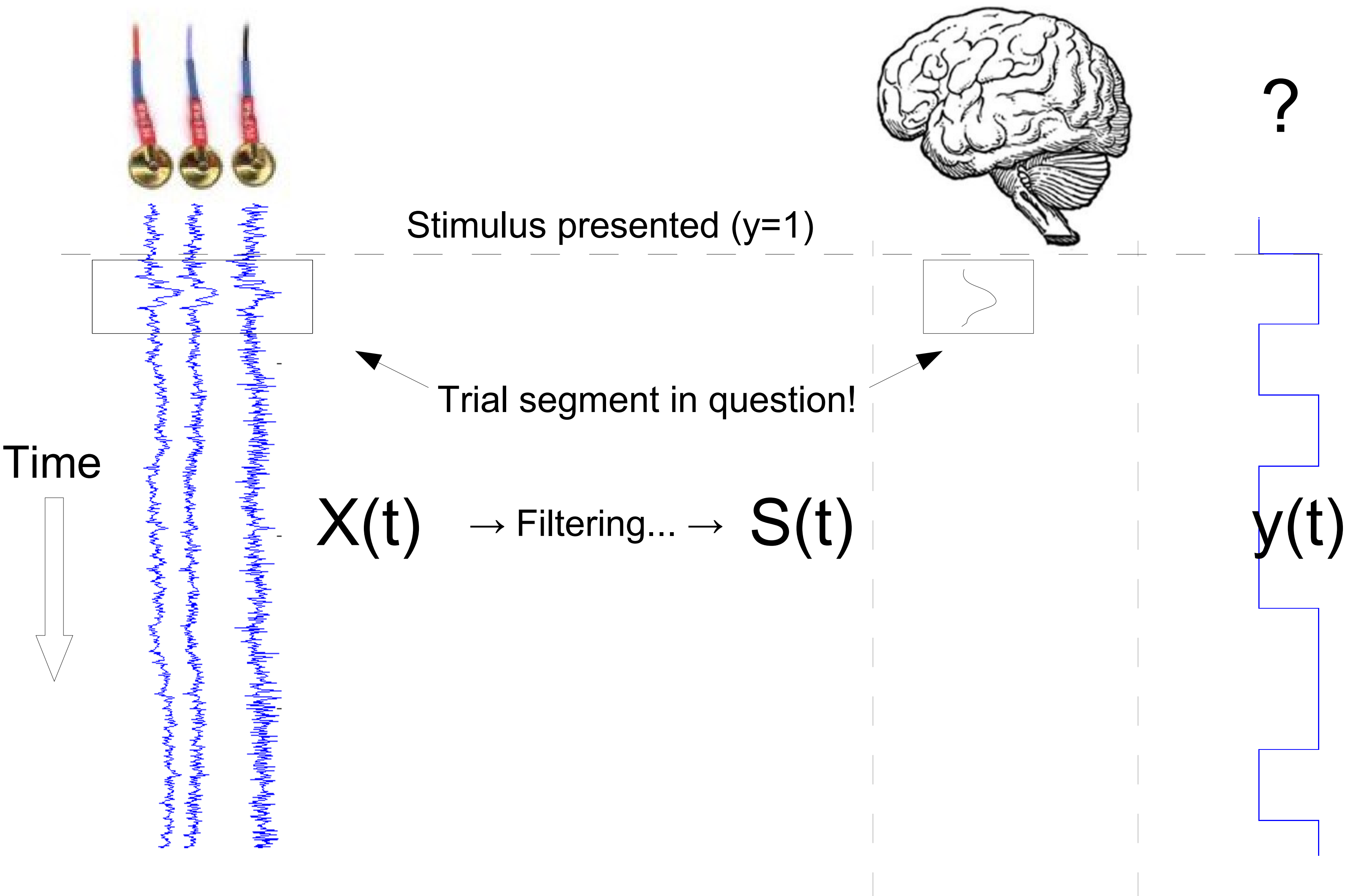
- A sequence of filters (e.g., spatial, spectral) is applied and each trial is extracted from the processed data



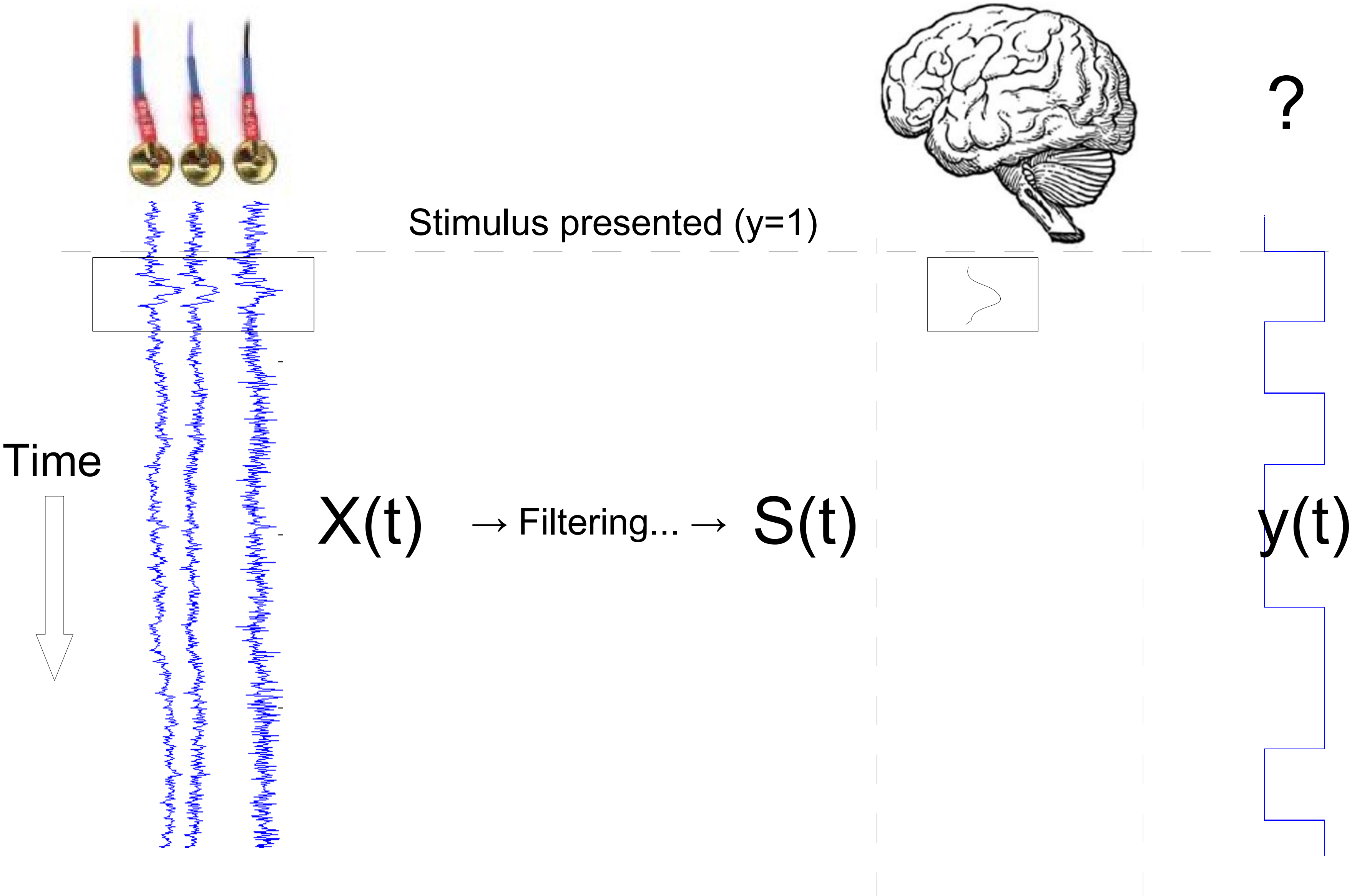
Concrete Example



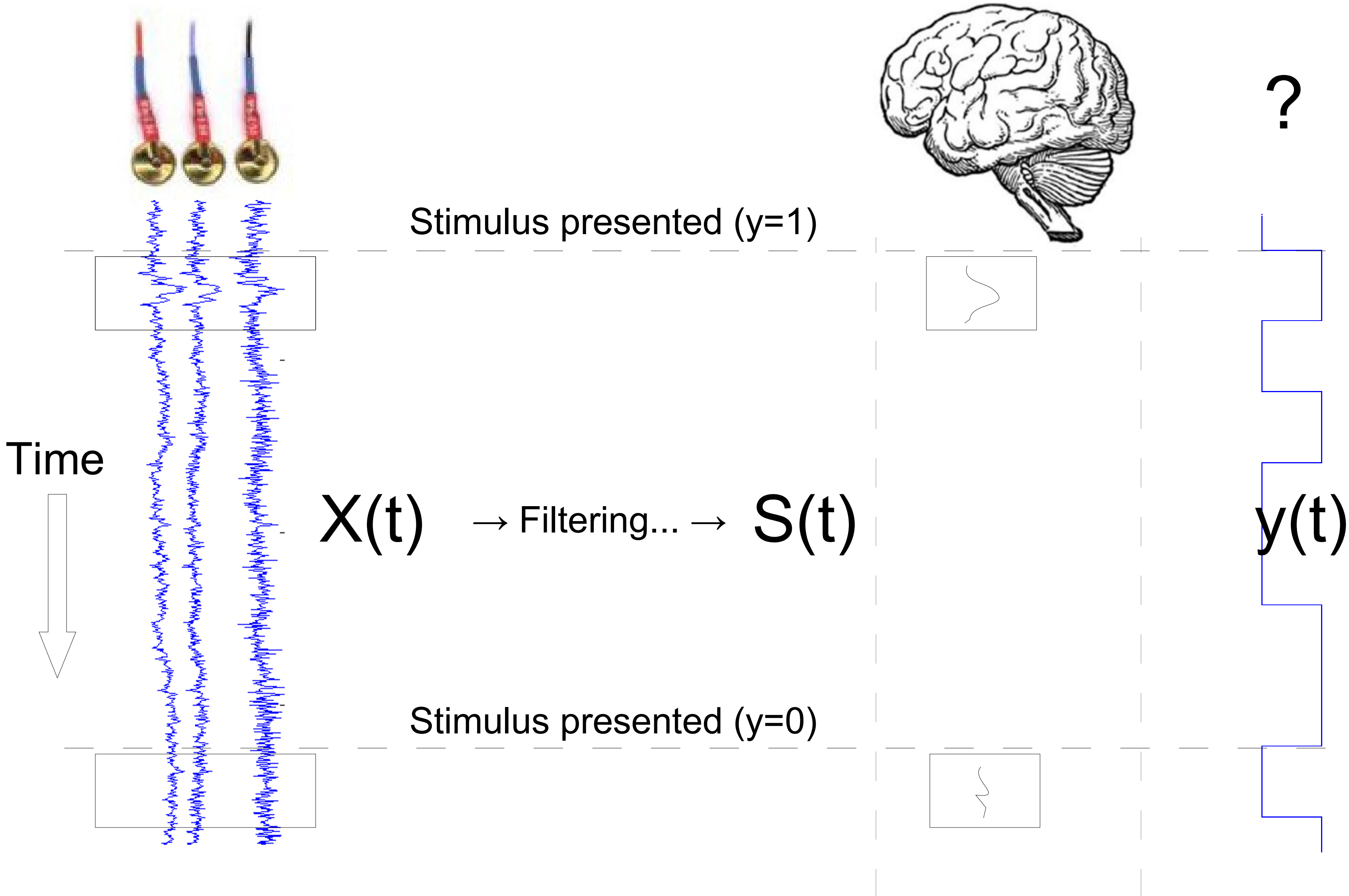
Concrete Example



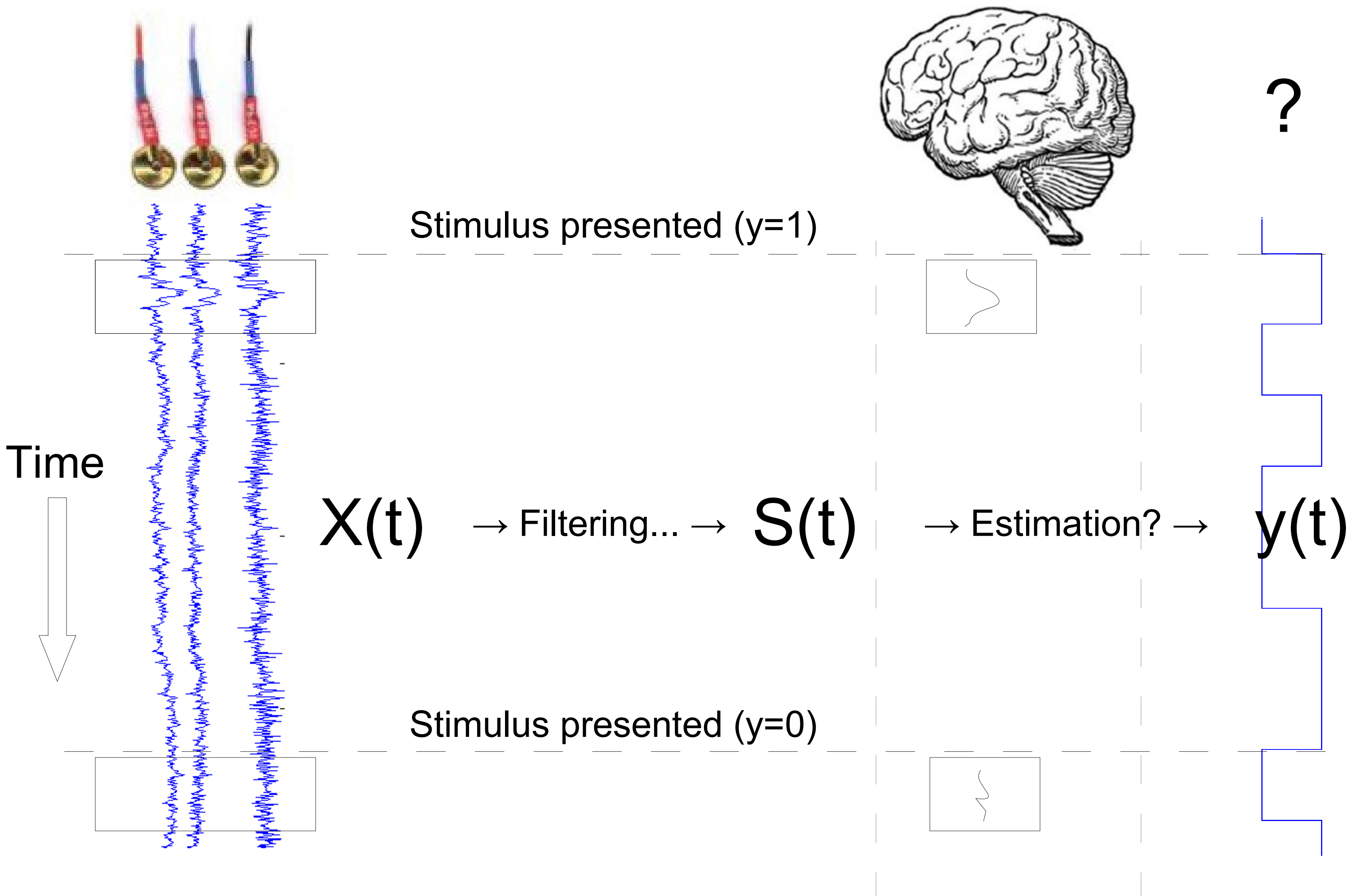
Concrete Example



Concrete Example

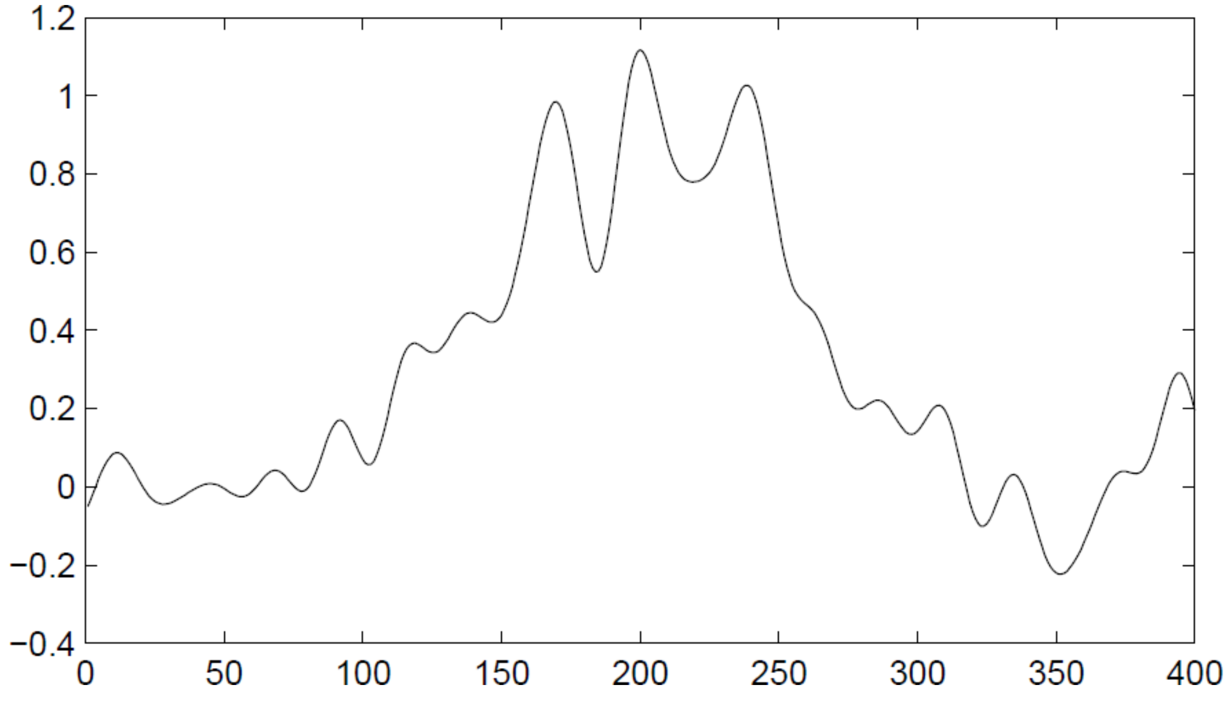


Concrete Example



Concrete Example

- For each resulting processed trial segment, a group of features that is assumed to carry information about the cognitive state at the given time interval is extracted

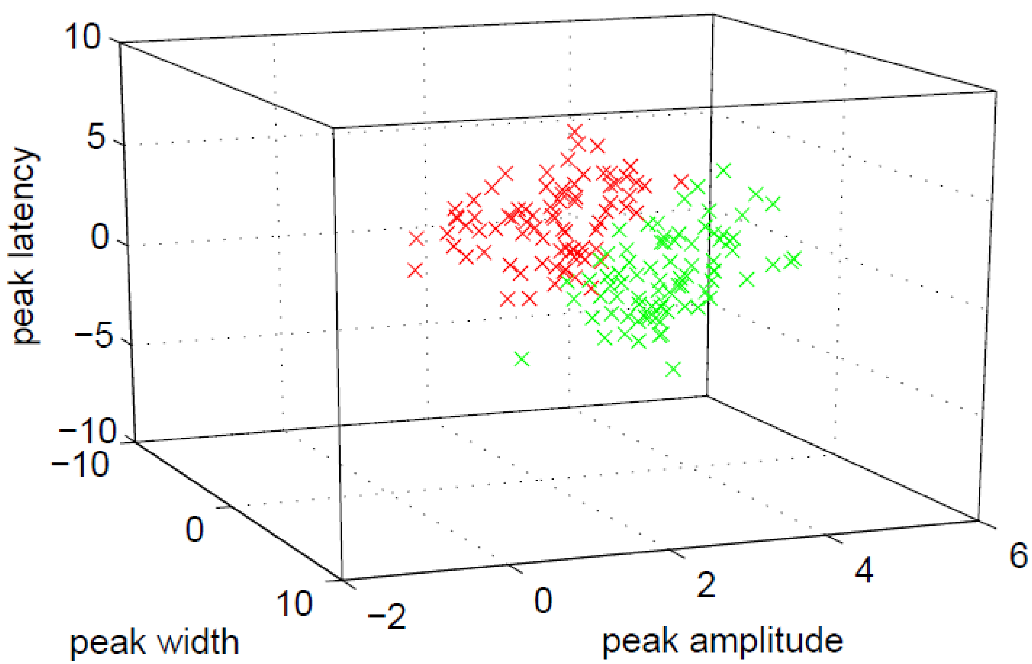


Concrete Example

- Feature extraction translates a processed trial into a feature vector, reducing the data dimensionality
- This is a very powerful approach to simplify the learning problem
- In this example, each trial is reduced to three parameters: peak latency, peak amplitude, peak width
- This implements the (very strong) assumption that a peak is the distinctive feature of each trial

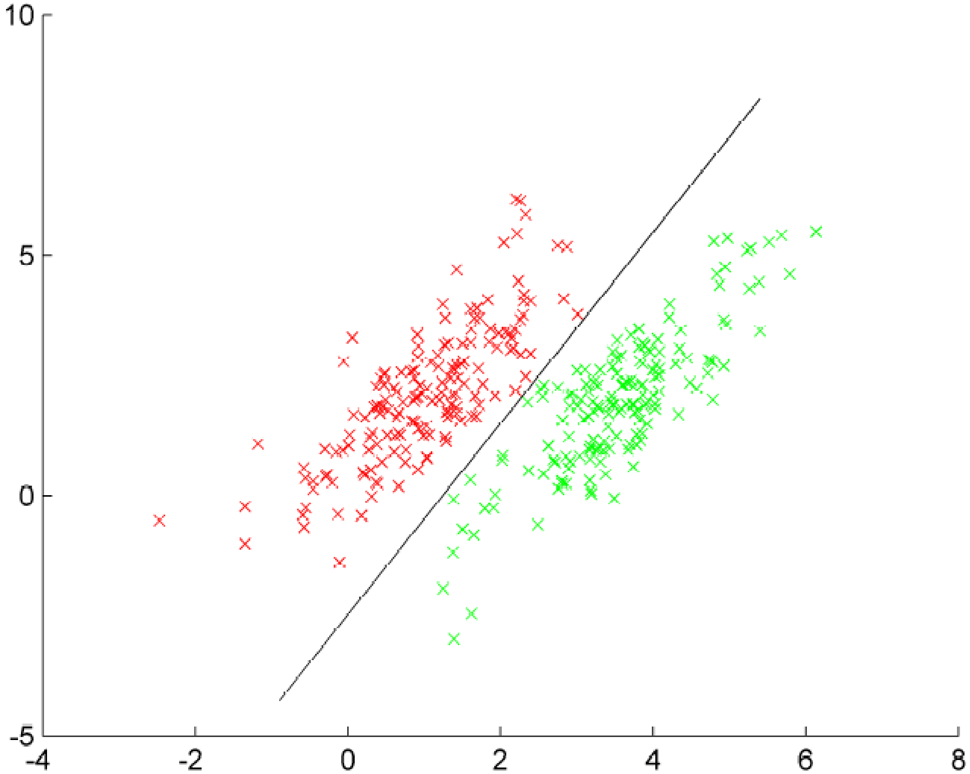
Concrete Example

- This gives a feature vector for every trial, which can be interpreted as the coordinates of a point in some space
- As the experimental condition for every trial is known, we learn the distribution of conditions in this feature space



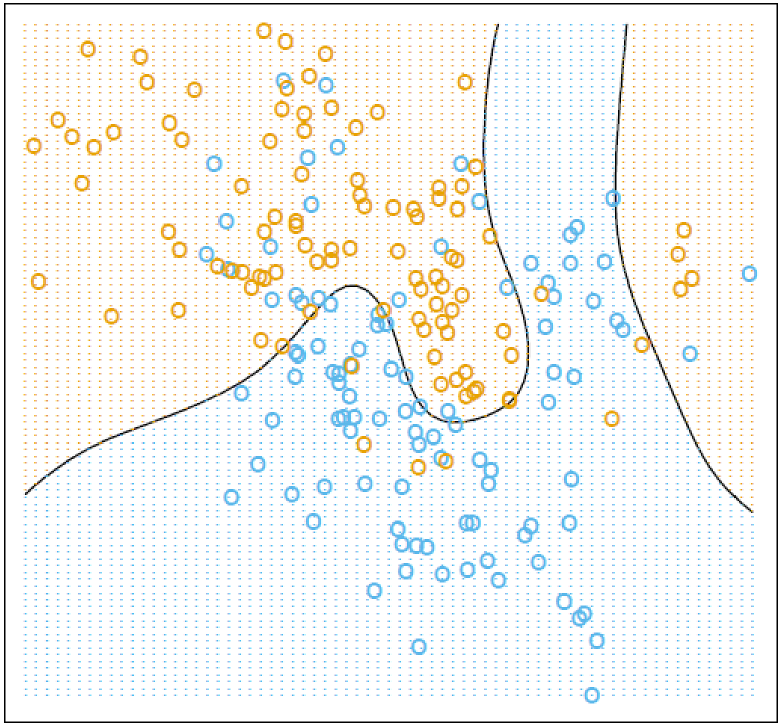
Concrete Example

- The task is to find a parametric mapping from locations in this space to final outputs
- The equation for a plane as below gives parameters w,b



Concrete Example

- In some cases, complex assumptions are required, as the data can be high-dimensional, overlapping, and can have complicated spatial structure

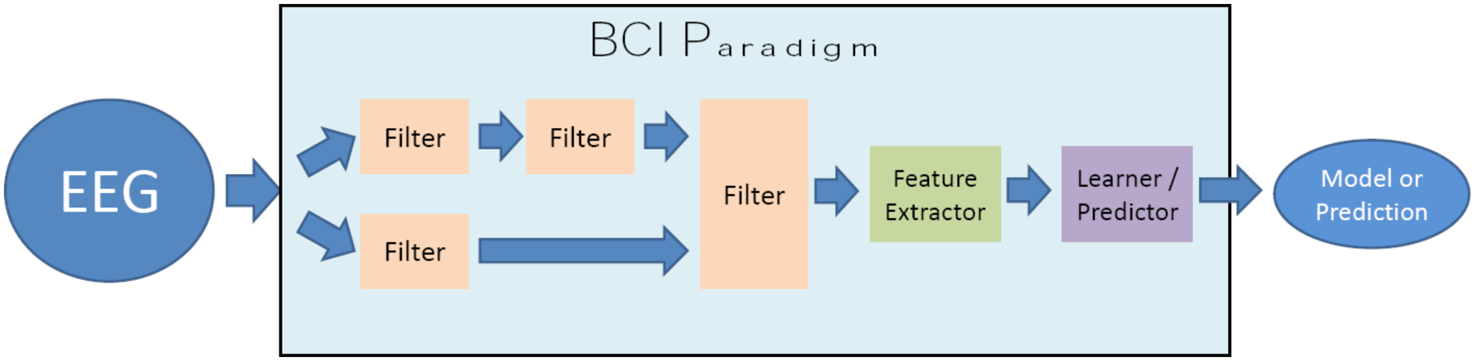


Concrete Example

- The result of this learning procedure is a predictive model (structure + parameters)
- This model can subsequently be
  - Applied in real time or “offline”
  - Evaluated on recorded data with known conditions
  - Inspected

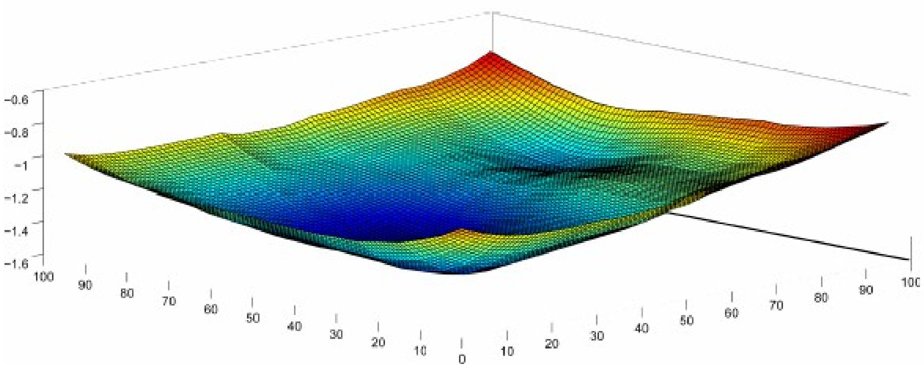
Typical BCI approach

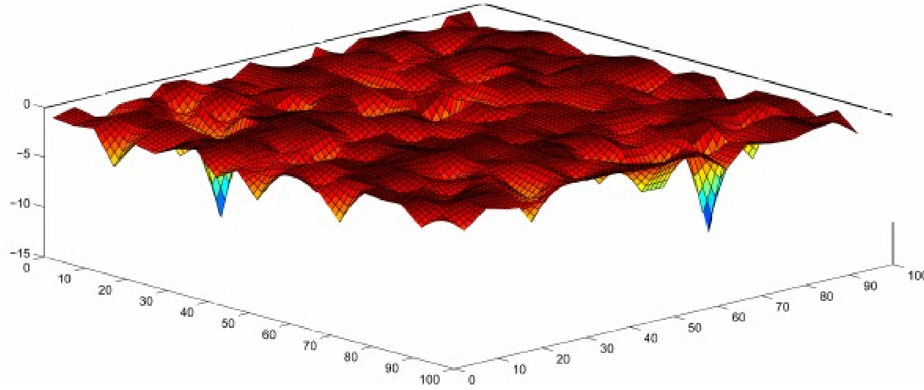
- An adaptive, learned, multi-stage parametric mapping (using signal processing and machine learning)
  - Preprocessing
  - Feature Extraction
  - Inference



Other possible approaches

- Solve all parameters as a joint optimization problem (can guarantee stronger optimality)



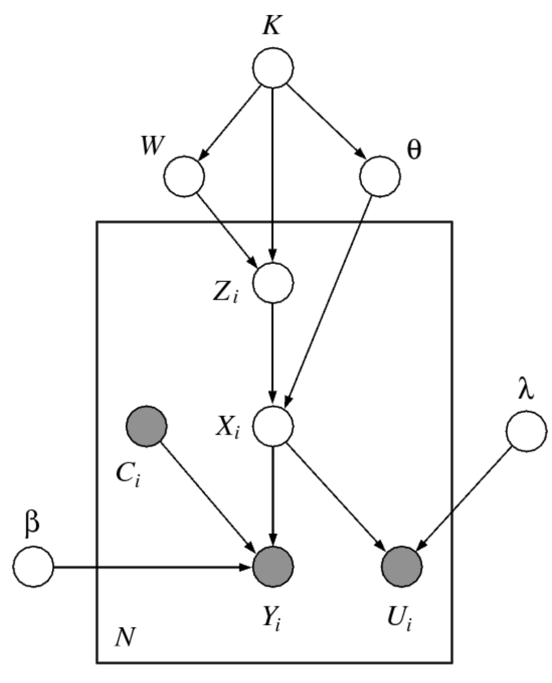


Easy

Hard!

Other possible approaches

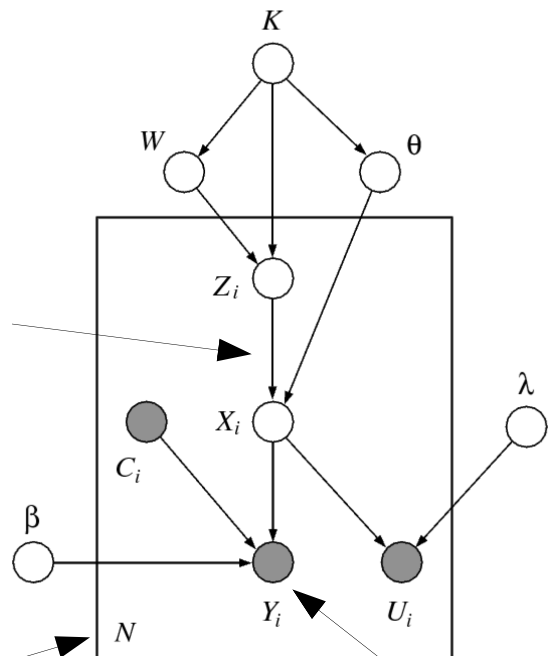
- Solve all parameters as a joint probabilistic inference problem (generalizes well to complex settings)



Jordan et al., 2005

Other possible approaches

- Solve all parameters as a joint probabilistic inference problem (generalizes well to complex settings)



Assumed latent structure (arrows)

Unobserved parameters (blank)

Replicated structure (e.g., N = #datasets)

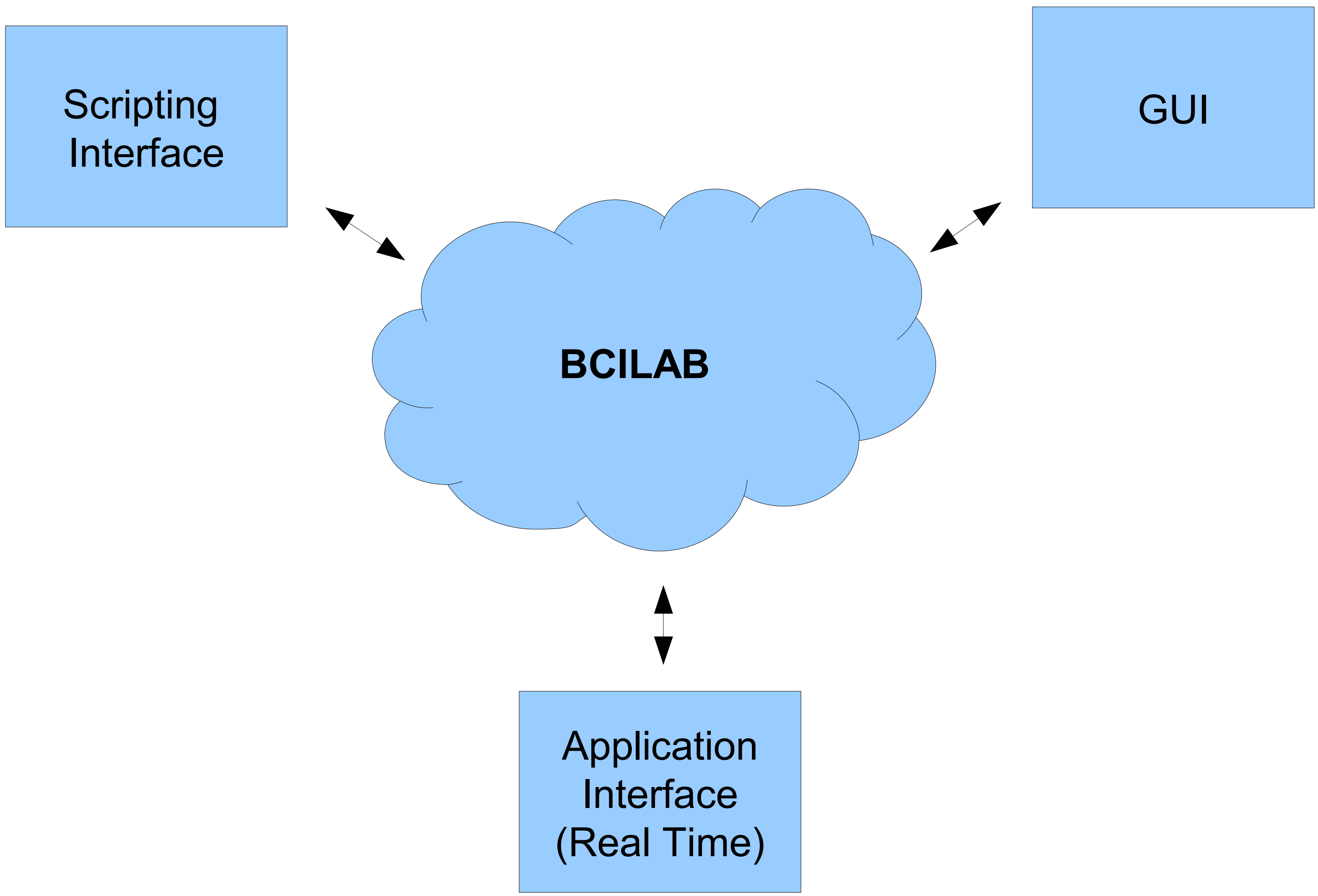
Observed parameters (shaded)

Jordan et al., 2005

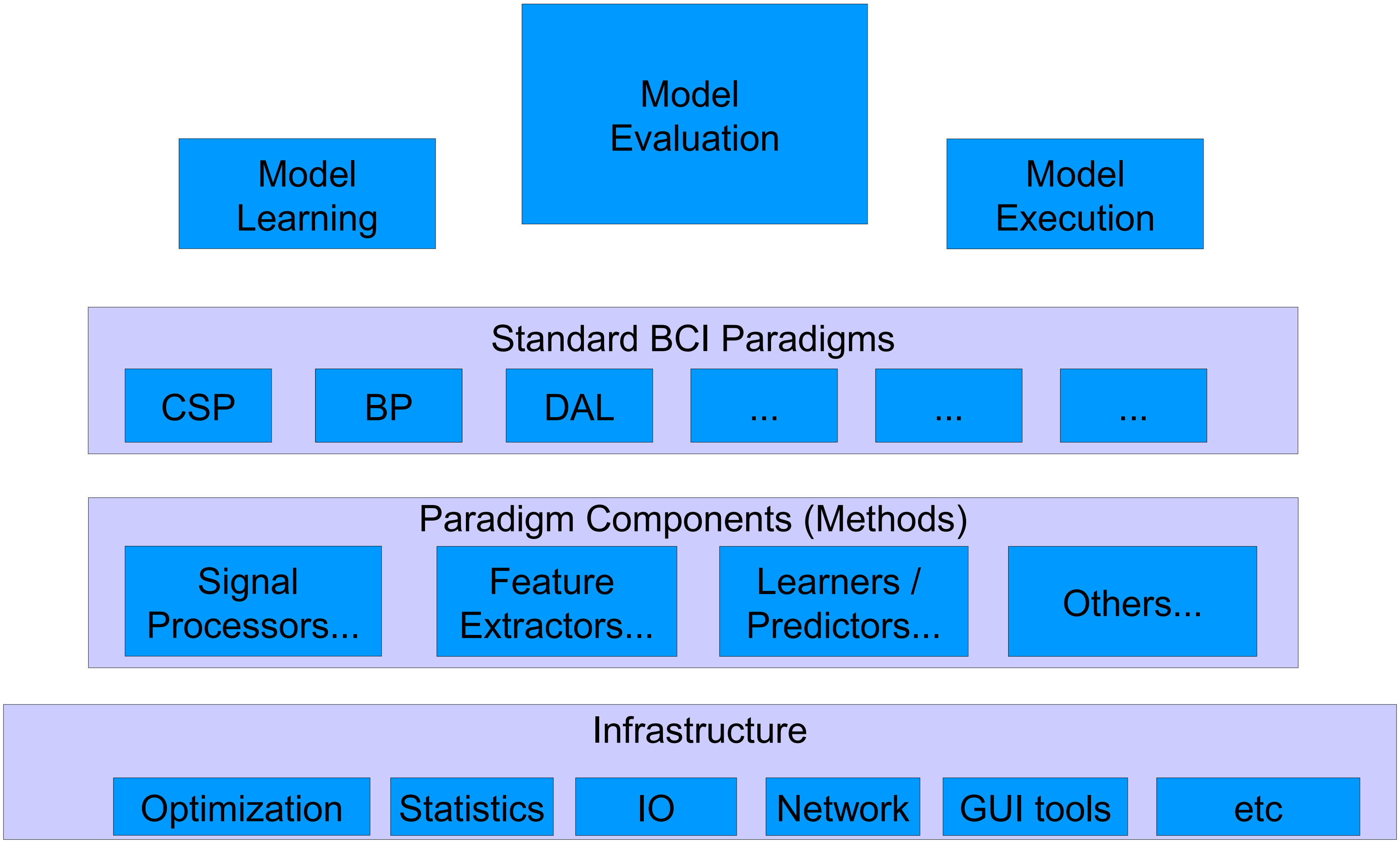
BCILAB

- BCI research is now supported by an EEGLAB plugin: the BCILAB toolbox
- It is a framework for the *design*, *calibration*, *evaluation* and *application* of Brain-Computer Interfaces
- The main goal is to offer the best achievable predictive performance, using state-of-the-art methods
- At the same time, it shall be accessible to non-experts, without restricting efficient use by experts

Toolbox Interfaces



Toolbox Structure



Thanks!

Questions?

(Next: practicum)