Machine Learning for Musical Interaction
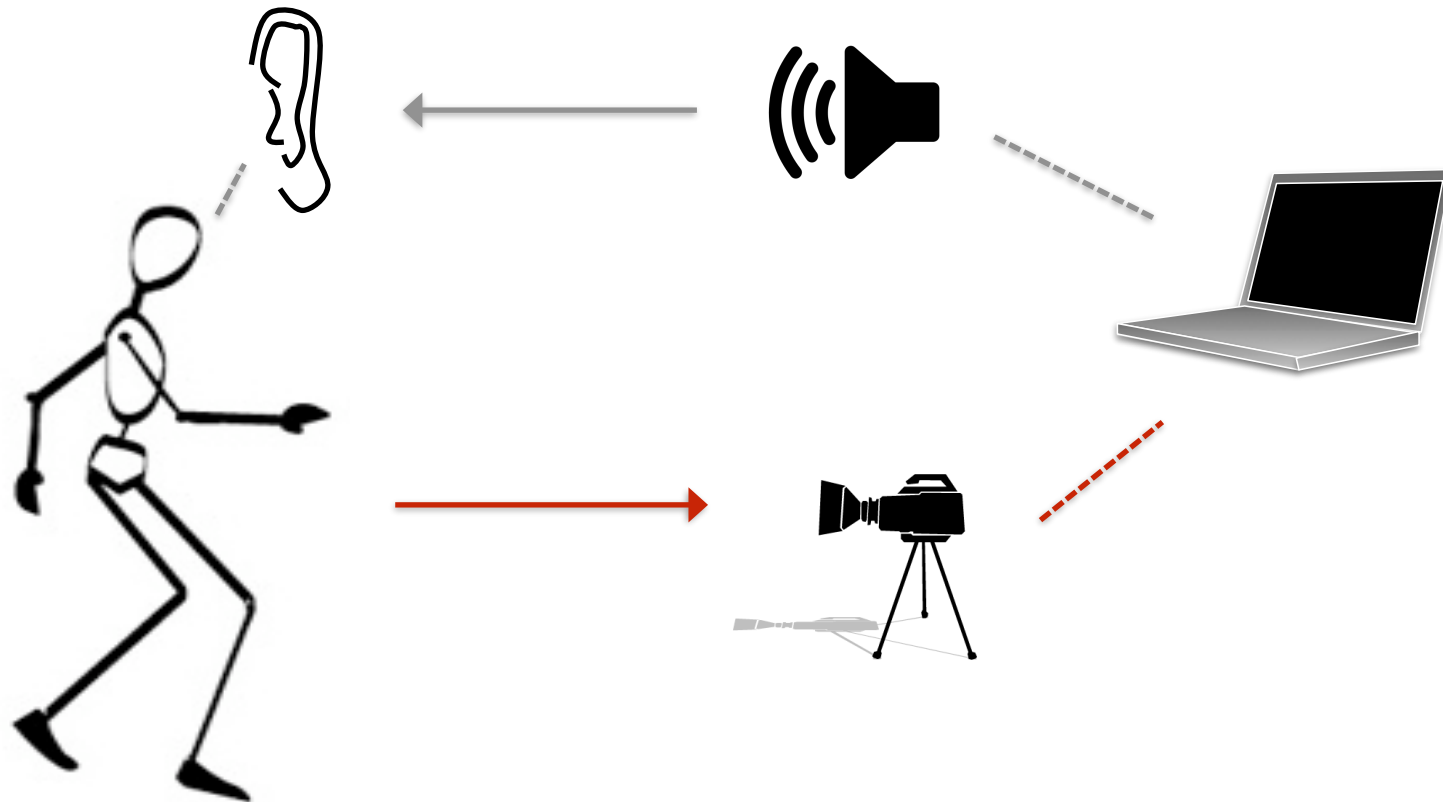
Baptiste Caramiaux
b.caramiaux@gold.ac.uk

EAVI group at Goldsmiths

- Embodied Audio-Visual Interaction
  http://eavi.goldsmithsdigital.com

- Machine Learning for Motion-based Musical Interaction

Problem(s)

How to write rules that would extract gesture characteristics from data?
How to write rules that would link gesture to sound automatically?

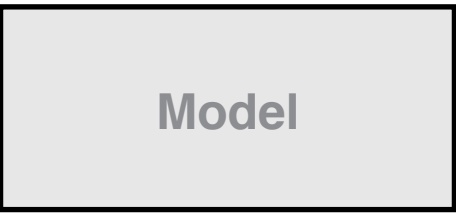Machine Learning (ML) is a body of **methods** that achieve **tasks** by **learning** from **examples**.
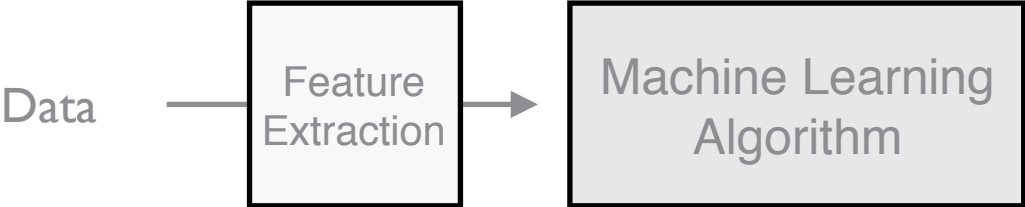
Step I
Learning

Data $\longrightarrow$

Machine Learning
Algorithm

Model

**Step I
Learning**

Data → Machine Learning Algorithm

**Step II
Acting**

New Data → Model → Action

**Step I
Learning**

Data — Feature Extraction → Machine Learning Algorithm

**Step II
Acting**

New Data — Feature Extraction → **Model** → Action
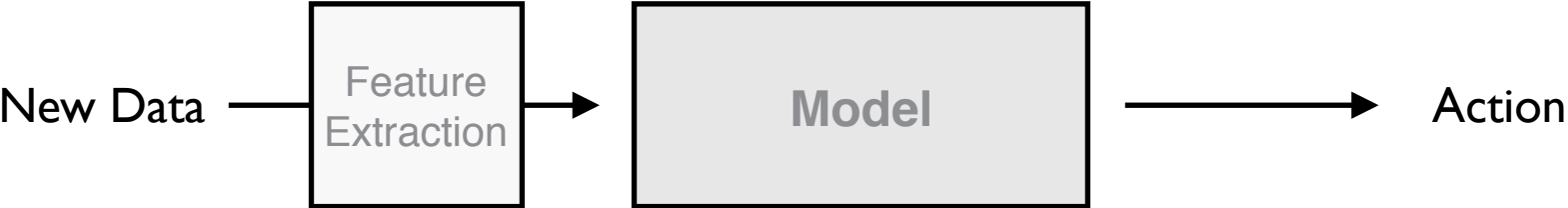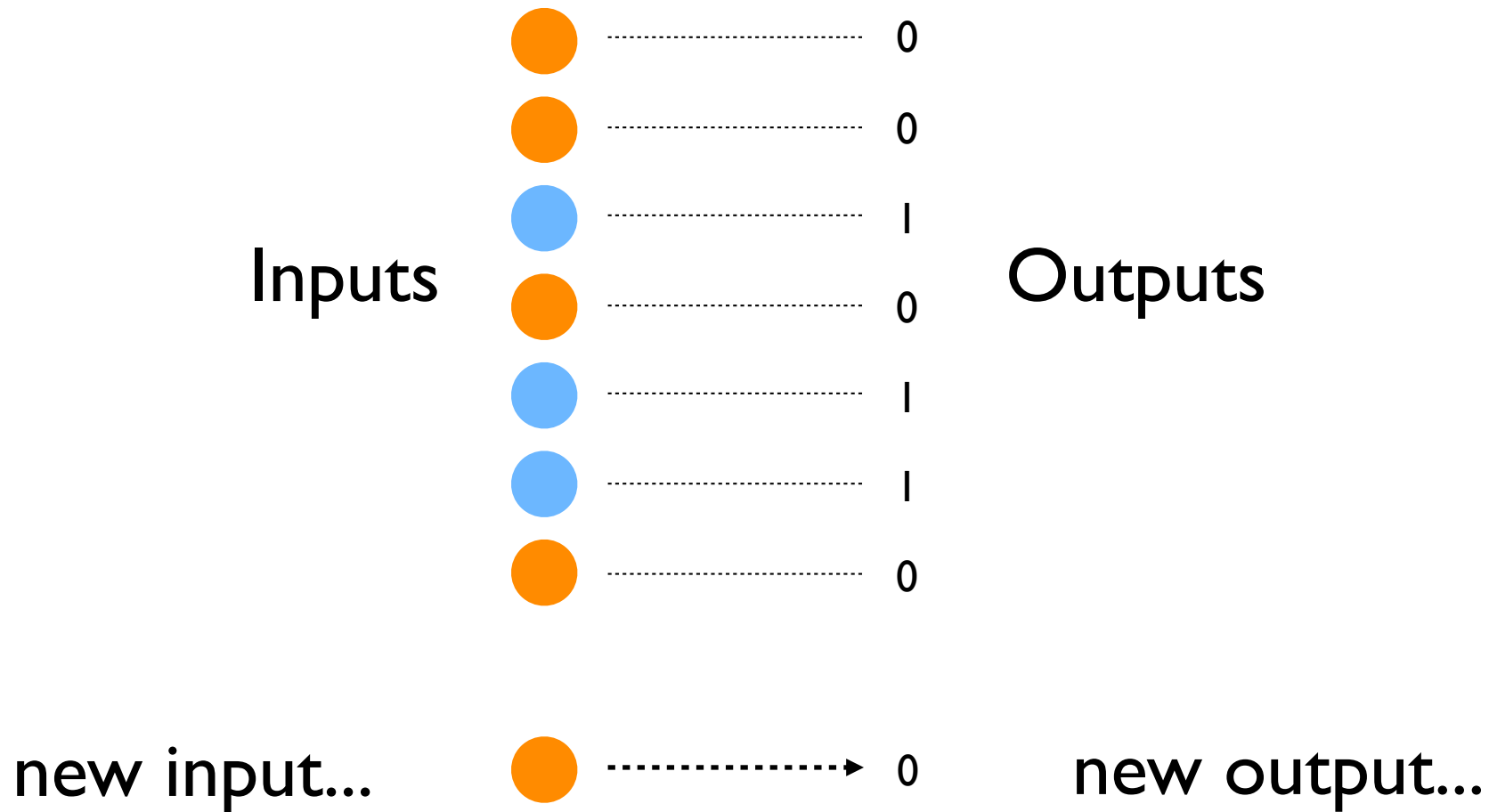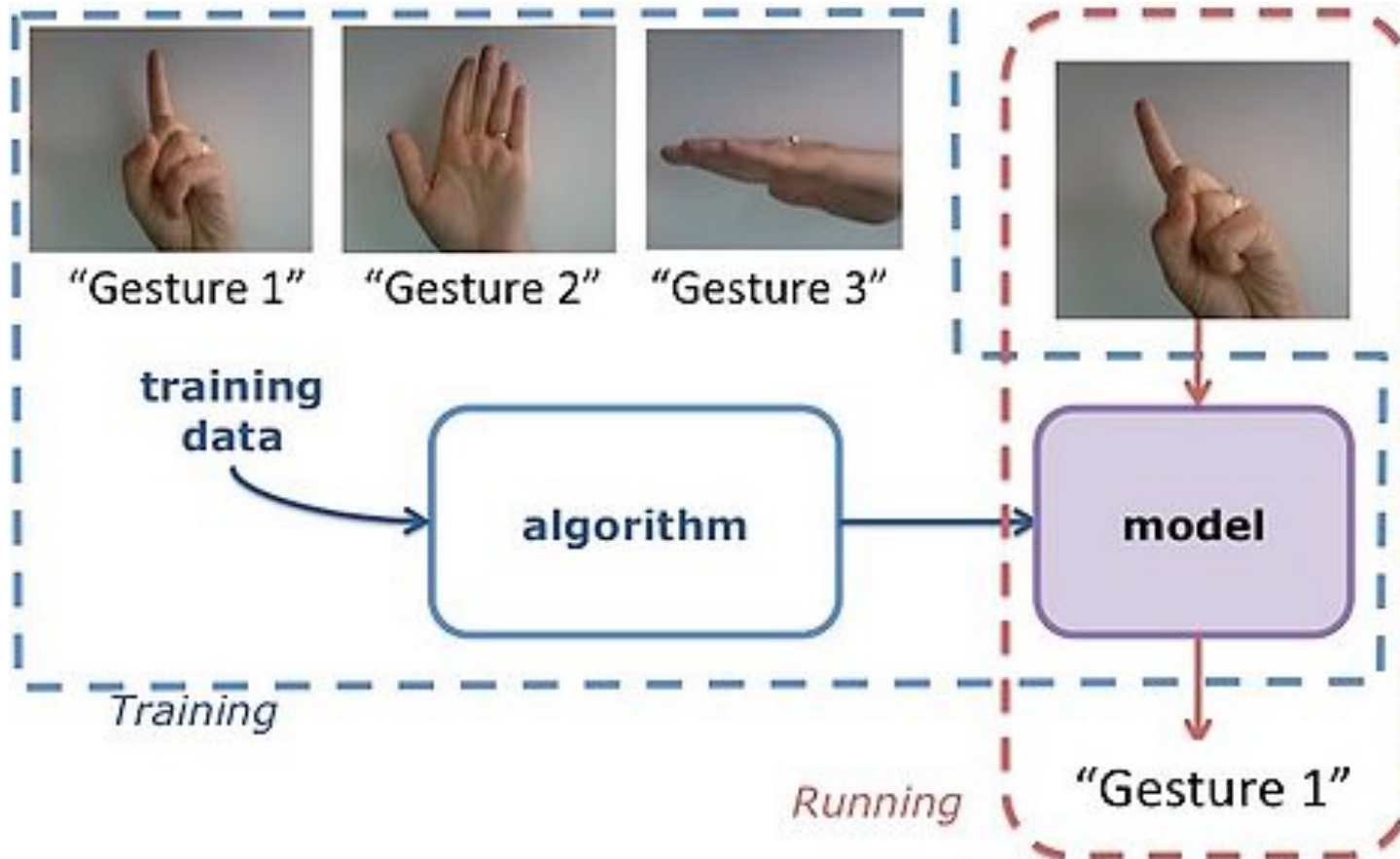
Learning from examples

- Finding relationships between elements under constraints

- Most used types of learning:

  - Supervised (examples are input–output)
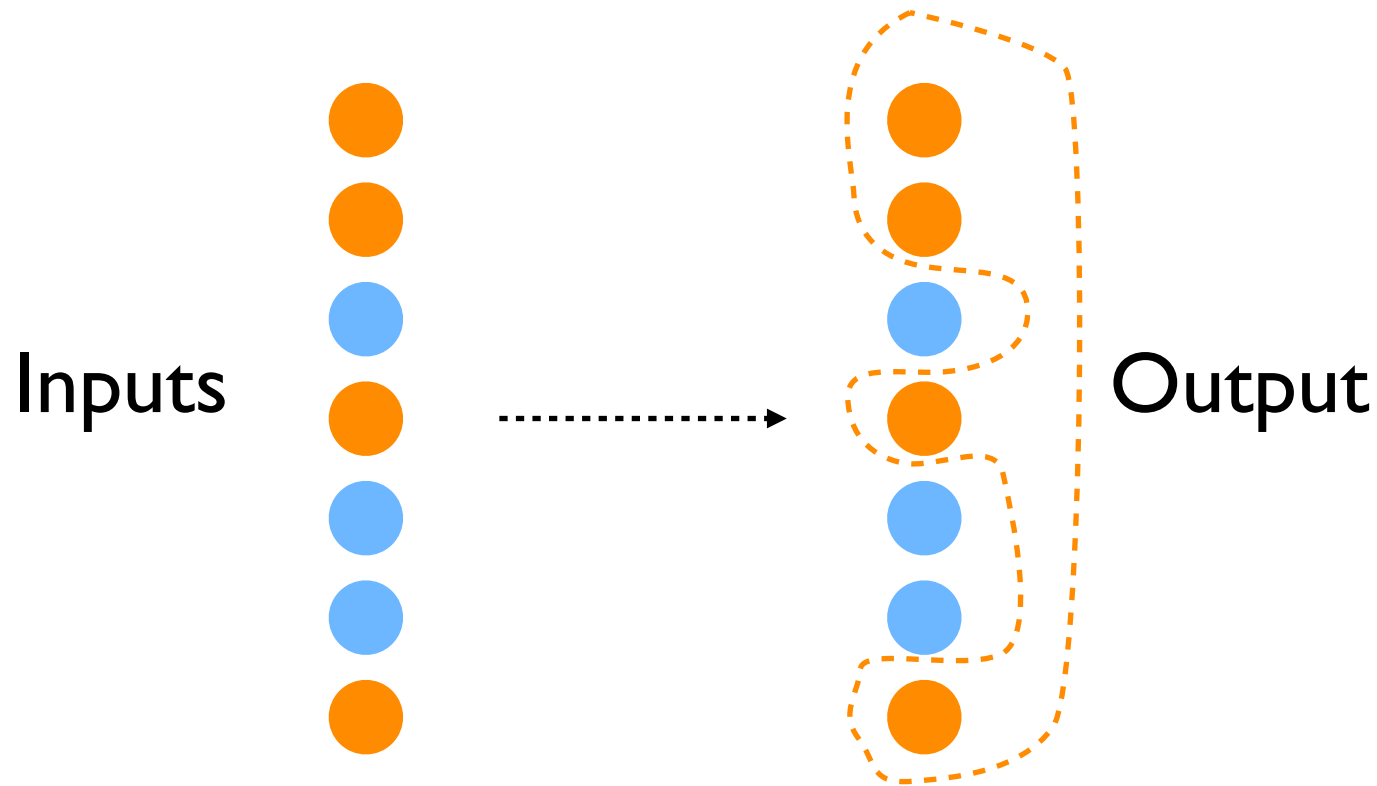
  - Unsupervised (examples are input only)

Supervised learning

Supervised learning

"Gesture 1"   "Gesture 2"   "Gesture 3"

training
data

algorithm

model

Training

Running

"Gesture 1"

22

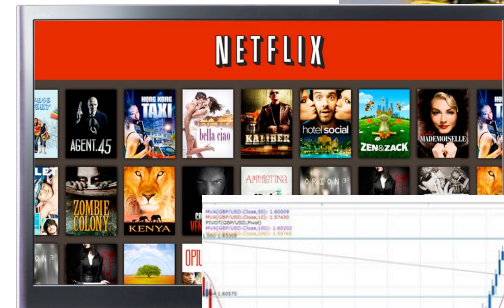*credits Rebecca Fiebrink*

Un-supervised learning

Inputs ....▸ Output

Applications

- Activity recognition

- Recommendation

- Finance forecasting
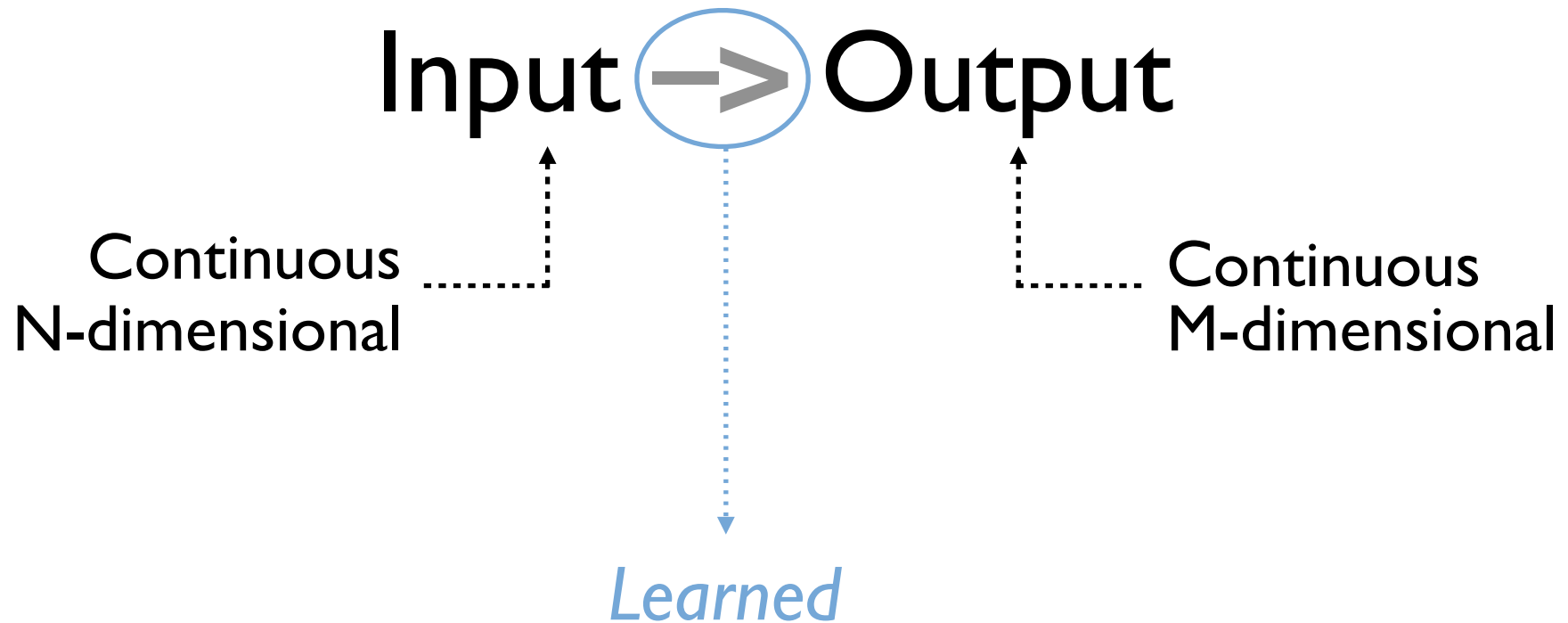
- ... and Musical interaction!

Machine Learning: Tasks

- Regression

- Classification

- Clustering

- Segmentation

- Forecasting...

Machine Learning: Tasks
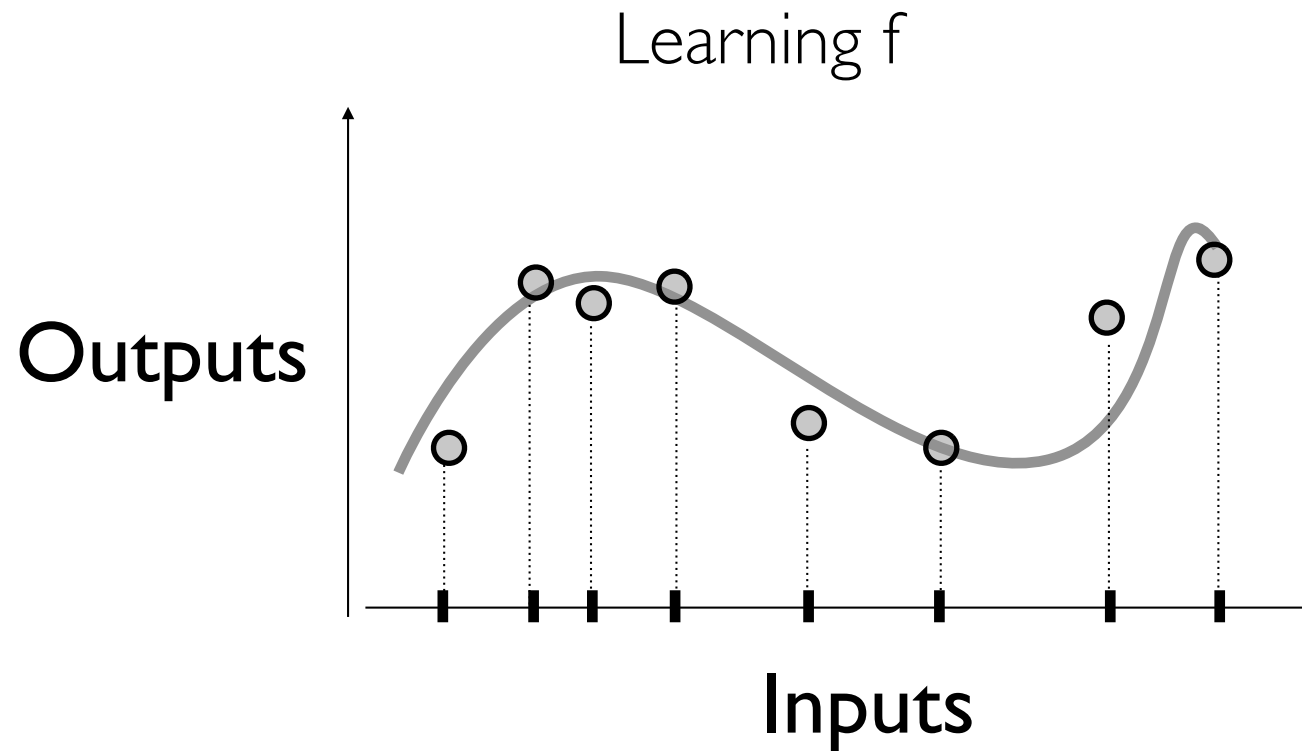
- Regression

- Classification

- Clustering

- Segmentation

- Forecasting...

Regression

Input ➔ Output

Continuous
N-dimensional

Continuous
M-dimensional

Learned

Learning

Example: Output = *f*(Input)

Learning f

Outputs

Inputs

Testing

Example: Output = *f*(Input)

Testing with a new input

Estimated output at *t*

Input at *t*
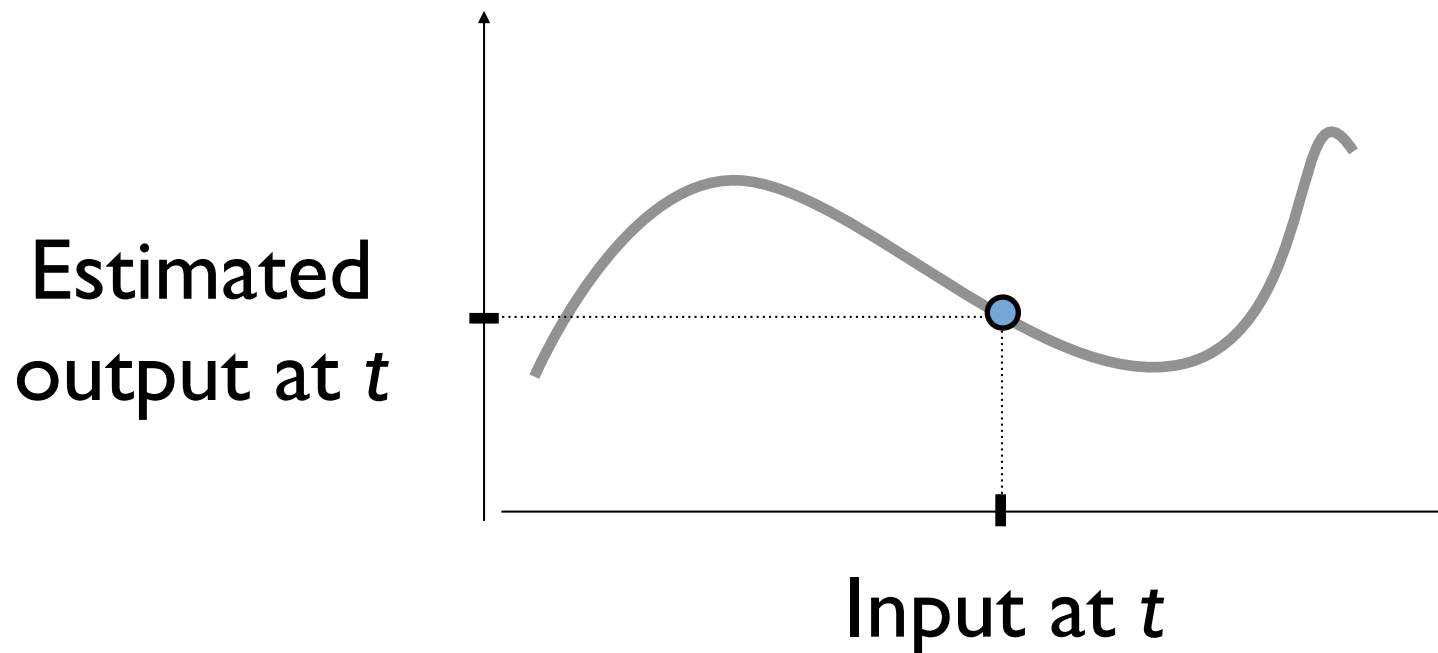
Demo

Applications

- Cross-modal control

- Gesture analysis

- Cross-modal analysis

Classification

Input ➔ Output (class)

Continuous
N-dimensional

Discrete

Learned
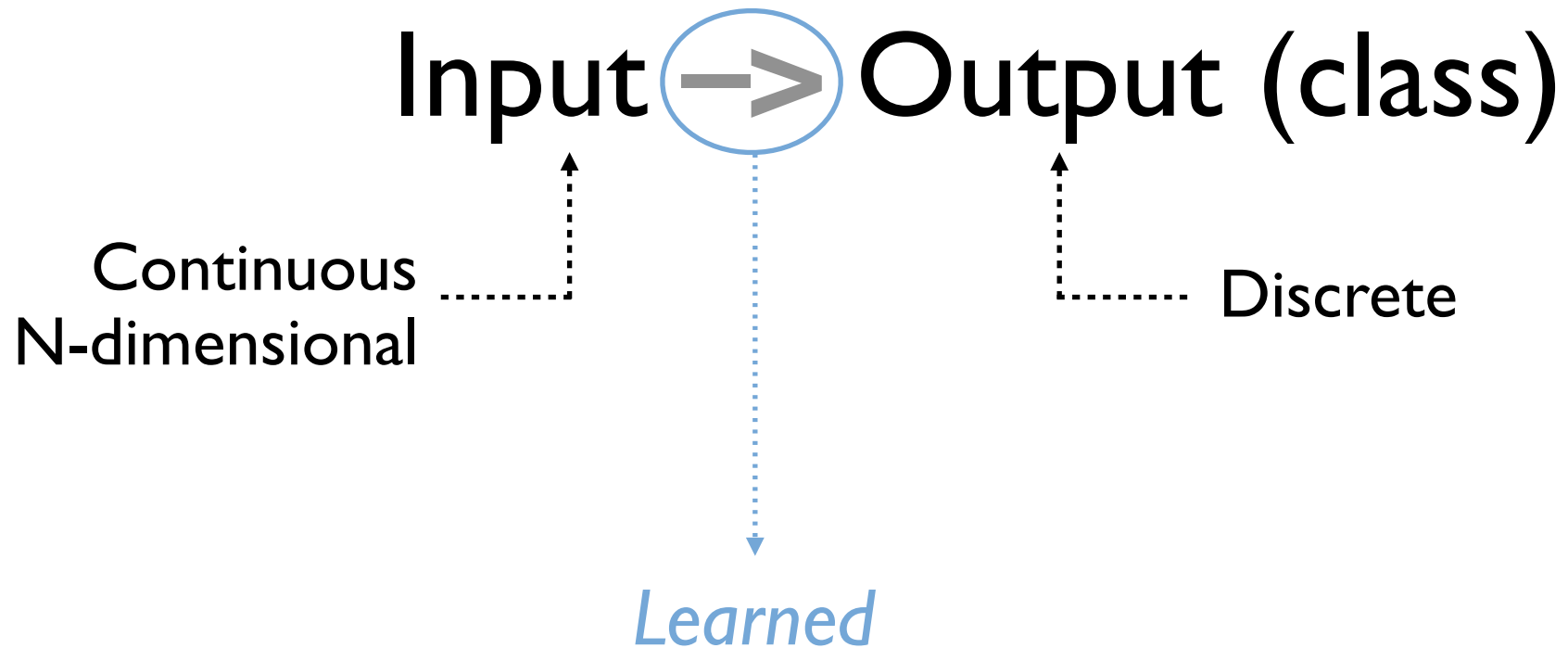
Learning

Example: Output = *f*(Input)

Learning f

Testing

Example: Output = *f*(Input)

Testing with a new input

Assigned class 0
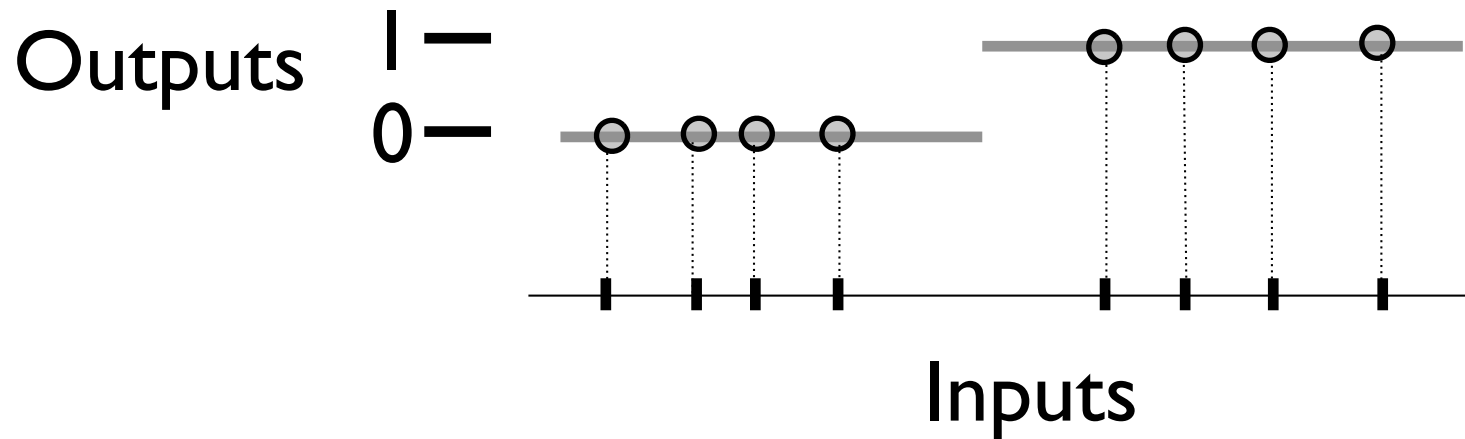
Input at *t*
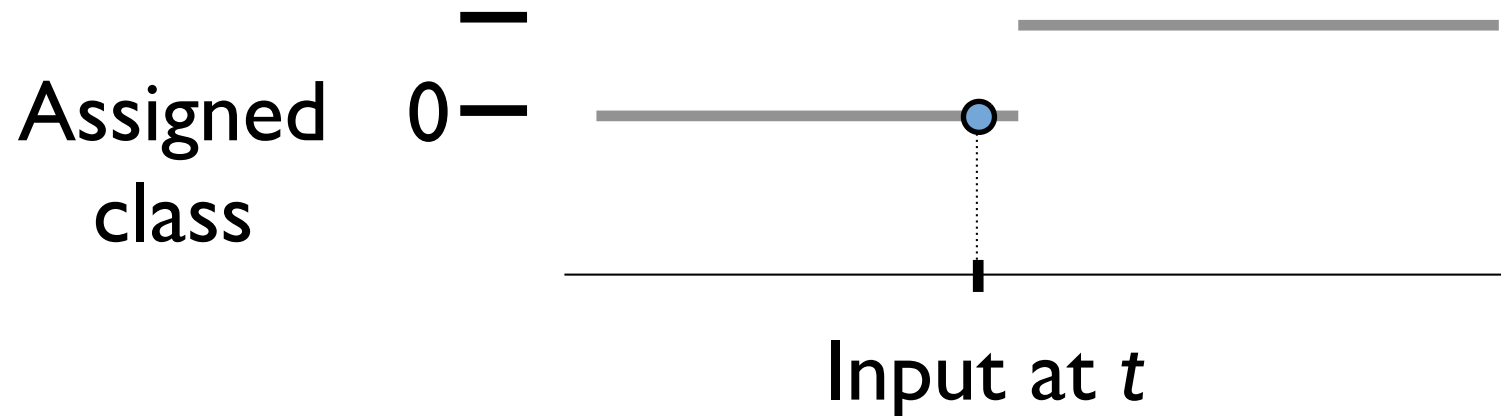
Demo

Examples

- Multi-parametric control
  C. Kiefer: http://vimeo.com/5173385
  S. Nicolls: http://vimeo.com/26678719

- Gesture analysis

- Gesture within time

Classification

Output = f(Input)

Class: 0

Input at t
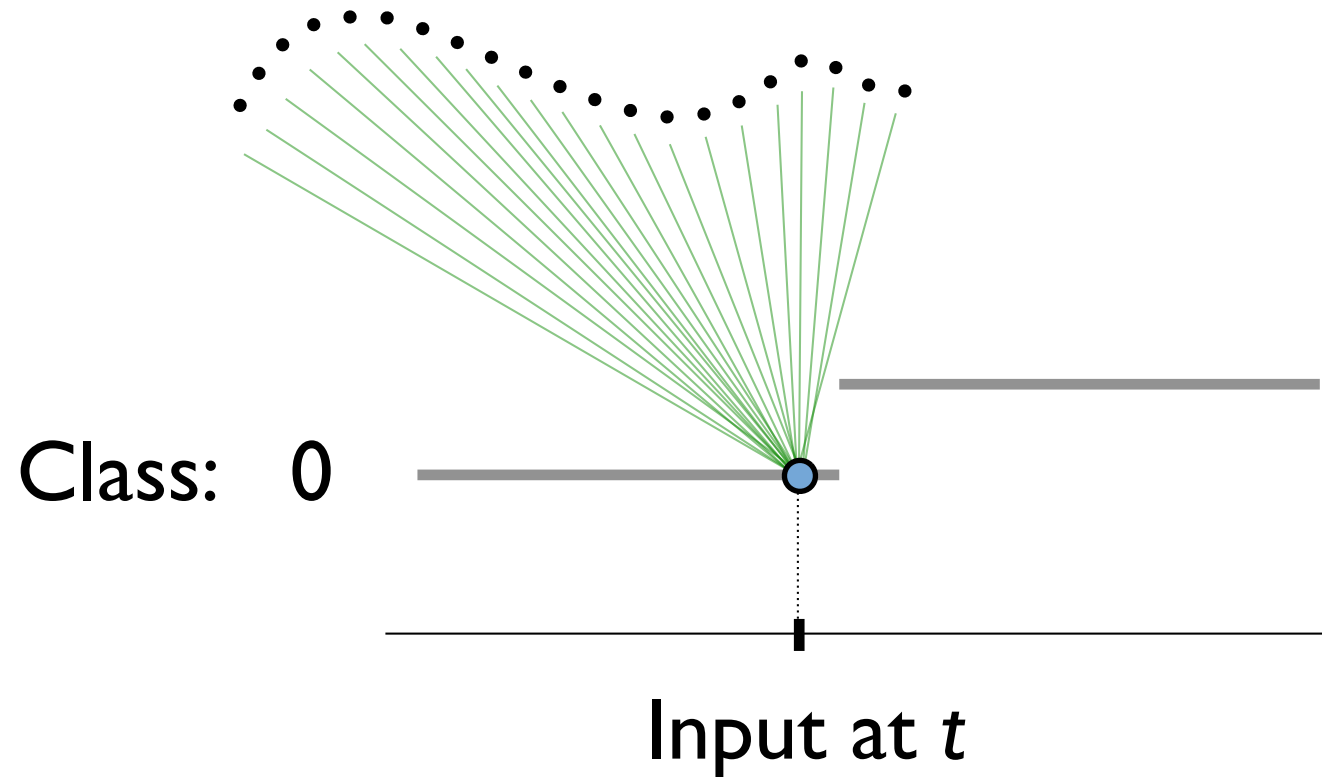
Realtime Recognition
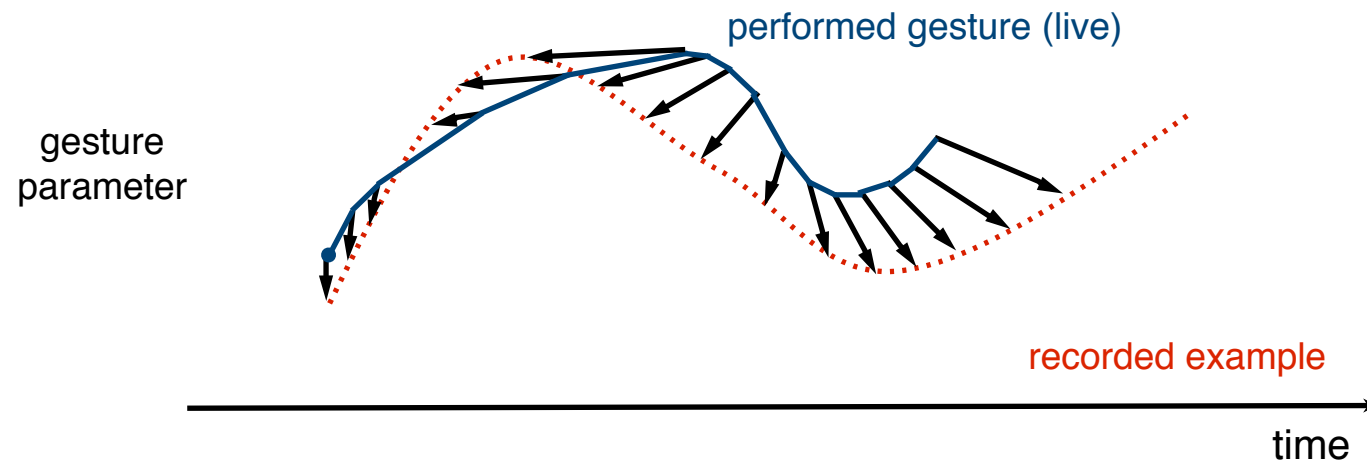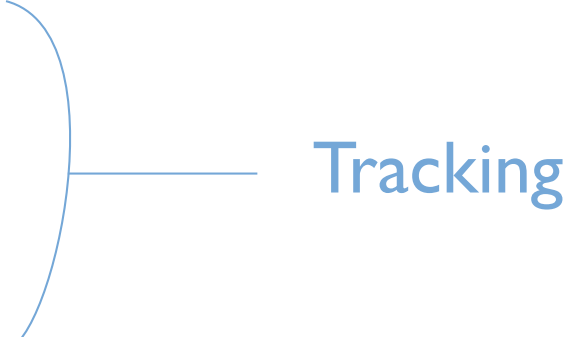
Gesture Variation Follower

Non-Linear Dynamical Systems + Particle Filtering
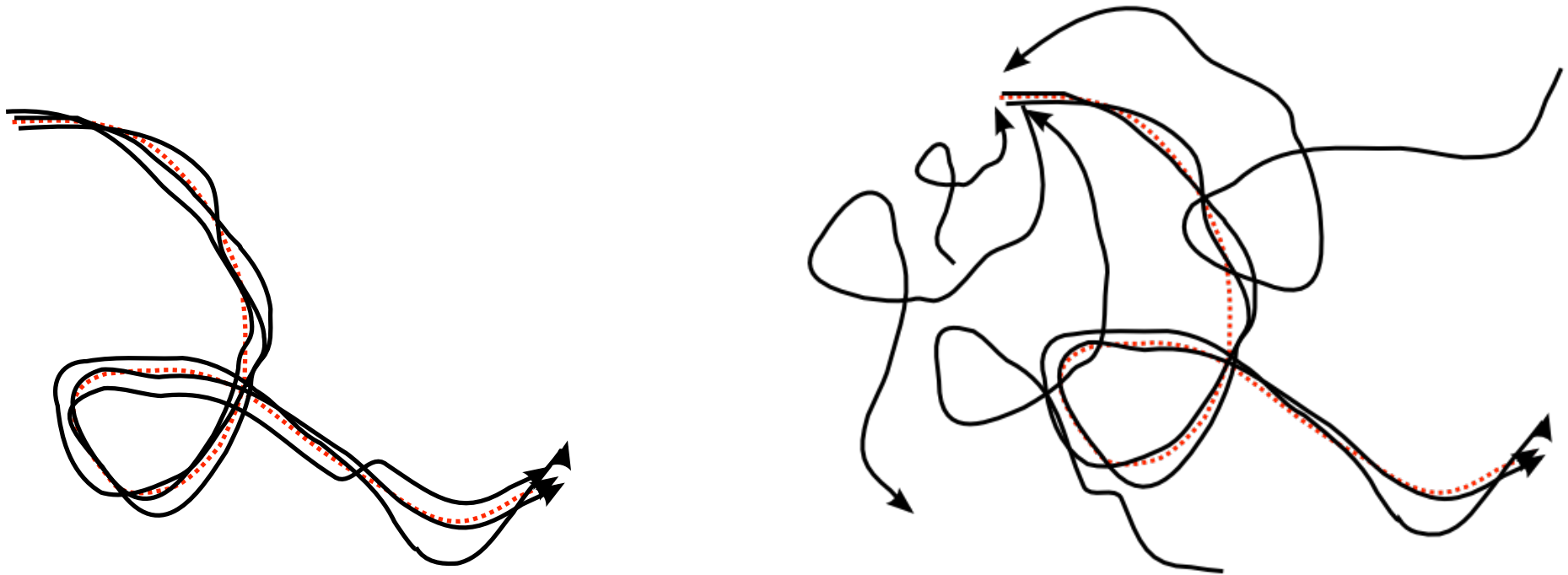
Gesture Variation Follower

Non-Linear Dynamical Systems + Particle Filtering

- Real-time gesture recognition   Classification

- Estimated features

  - Time progression
  - Speed   Tracking
  - Others: scale, rotation, ...

Demos

Research

- Interaction Design

- Usability

- User experience

- Performances

Summary

- Machine learning techniques learn from examples

- What technique should I use?

  - Scenario building

    ▶ Learning method: supervised / unsupervised

    ▶ Task: classification / regression

    ▶ Information to be considered: temporal structure, variations, etc.

  - Iterations...

Practice

Gestural Sound Toolkit

Gestural Sound Toolkit

https://github.com/bcaramiaux/Gestural-Sound-Toolkit

Gestural Sound Toolkit
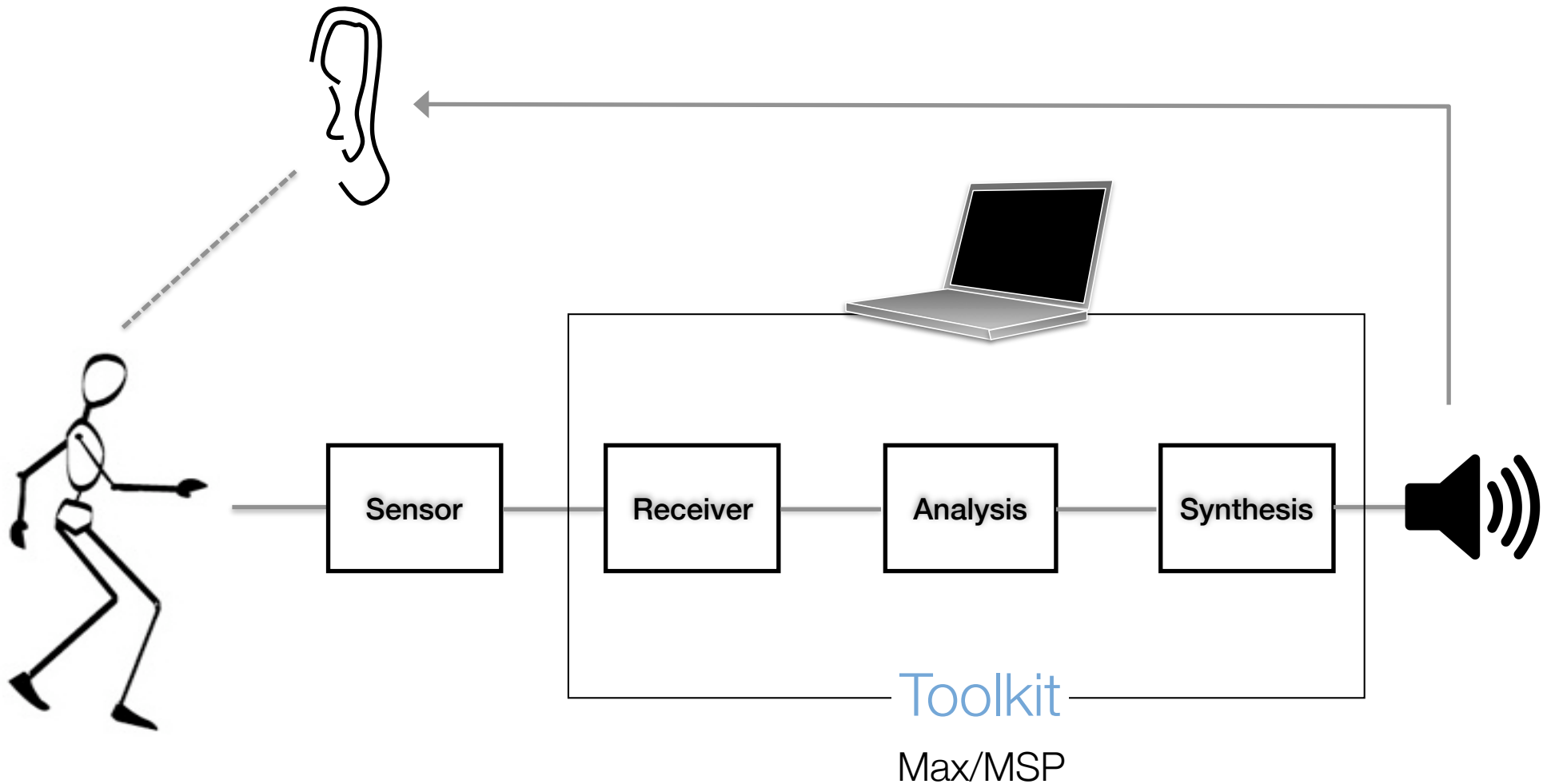
https://github.com/bcaramiaux/Gestural-Sound-Toolkit

b.caramiaux@gold.ac.uk

Installation

- Get from:
  https://github.com/bcaramiaux/Gestural-Sound-Toolkit

- Open Max/MSP

- In Max: Options -> File Preference

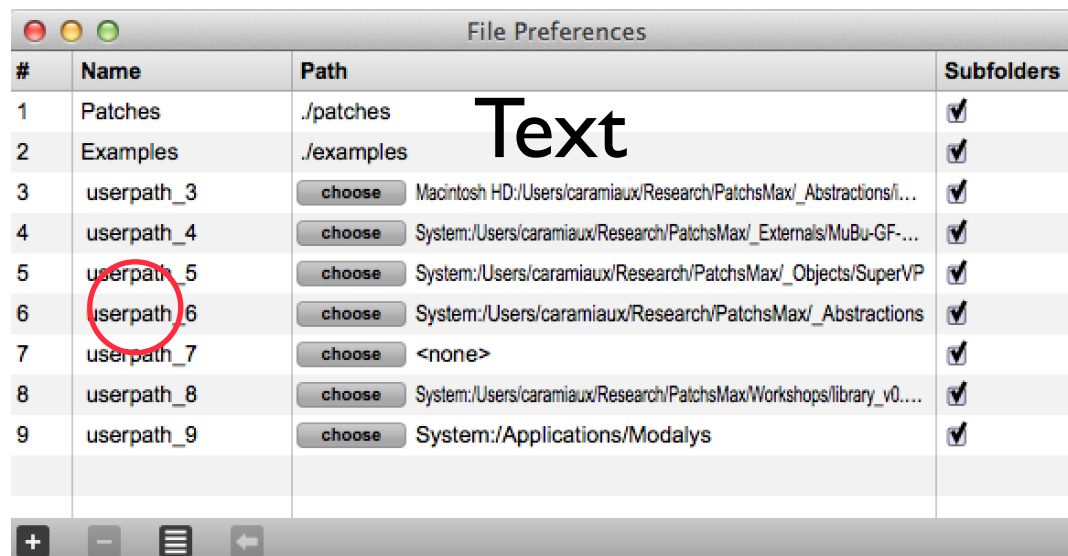| #	Name	Path		Subfolders
1	Patches	./patches	Text	☑
2	Examples	./examples		☑
3	userpath_3	choose	Macintosh HD:/Users/caramiaux/Research/PatchsMax/_Abstractions/i...	☑
4	userpath_4	choose	System:/Users/caramiaux/Research/PatchsMax/_Externals/MuBu-GF-...	☑
5	userpath_5	choose	System:/Users/caramiaux/Research/PatchsMax/_Objects/SuperVP	☑
6	userpath_6	choose	System:/Users/caramiaux/Research/PatchsMax/_Abstractions	☑
7	userpath_7	choose	<none>	☑
8	userpath_8	choose	System:/Users/caramiaux/Research/PatchsMax/Workshops/library_v0....	☑
9	userpath_9	choose	System:/Applications/Modalys	☑

"choose" the folder "Gestural-Sound-Toolkit"

- Open "overview.maxpat" from "Gestural-Sound-Toolkit"