Structured stochastic processes and functional data analysis for the assessment of motor learning in normal and pathological subjects

(Postdoctoral research submitted to FAPESP)

Noslen Hernández

Antonio Galves, Claudia Vargas

II NeuroMat Workshop: New Frontiers in Neuromathematics

November 2016

Recent results [A. Duarte, R. Fraiman, A. Galves, G. Ost and C.D. Vargas, arXiv:1602.00579]

- Hypothesis: The brain retrieves statistical regularities from stimuli

Recent results [A. Duarte, R. Fraiman, A. Galves, G. Ost and C.D. Vargas, arXiv:1602.00579]

- Hypothesis: The brain retrieves statistical regularities from stimuli

- Definition of new kind stochastic processes

Recent results [A. Duarte, R. Fraiman, A. Galves, G. Ost and C.D. Vargas, arXiv:1602.00579]

- Hypothesis: The brain retrieves statistical regularities from stimuli

- Definition of new kind stochastic processes

Stochastic process driven by context tree model

Recent results [A. Duarte, R. Fraiman, A. Galves, G. Ost and C.D. Vargas, arXiv:1602.00579]

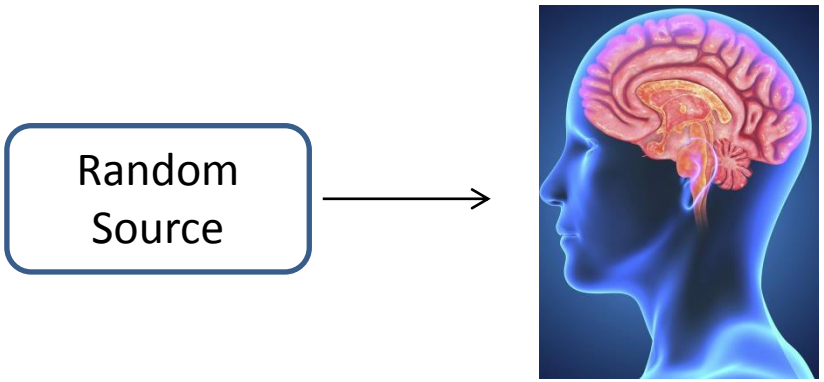
- Hypothesis: The brain retrieves statistical regularities from stimuli

- Definition of new kind stochastic processes

  Stochastic process driven by context tree model

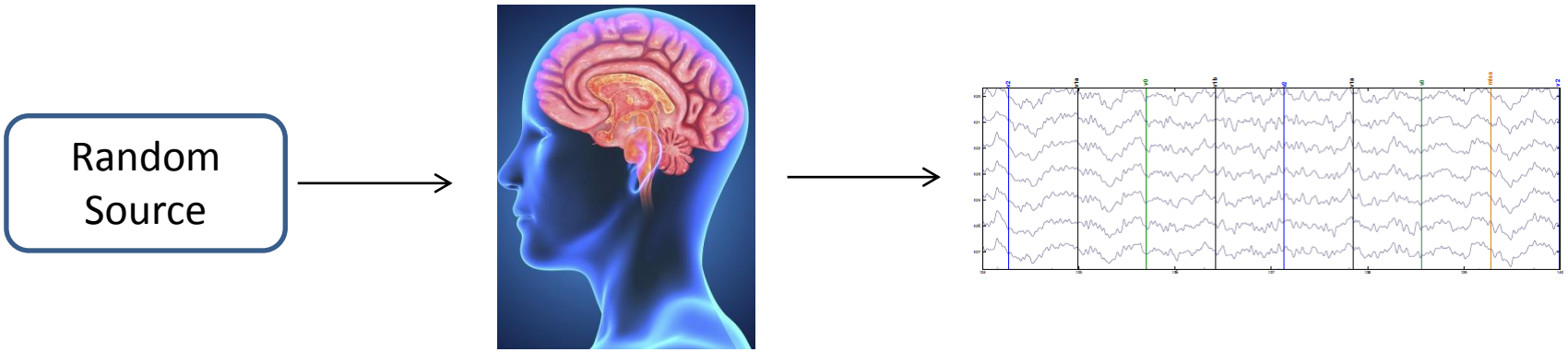
- Allows a design, modeling and analysis of neurophysiological experiments with structured stimuli

Retrieving a context tree form EEG data



Images taken from [A. Duarte, R. Fraiman, A. Galves, G. Ost and C.D. Vargas, arXiv:1602.00579]

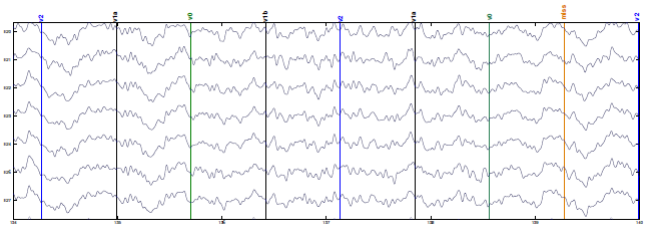
Retrieving a context tree form EEG data



Images taken from [A. Duarte, R. Fraiman, A. Galves, G. Ost and C.D. Vargas, arXiv:1602.00579]

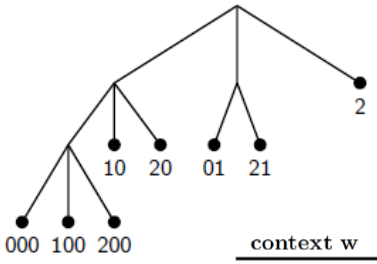
Retrieving a context tree form EEG data

Random Source



$\cdots 2\ 1\ 0\ 1\ 2\ 1\ 0\ 1\ 2\ 1\ 0\ 1\ 2 \cdots$

$\cdots 2\ 1\ 0\ 0\ 2\ 1\ 0\ 1\ 2\ 0\ 0\ 0\ 2 \cdots$

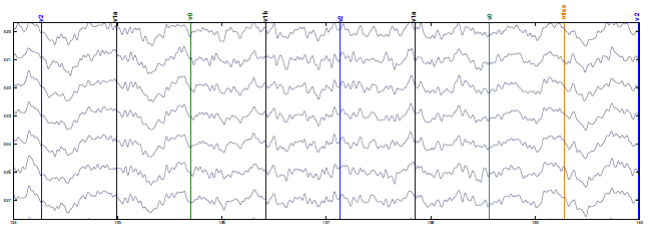


context w	p(1|w)	p(2|w)
2	$1-\epsilon$	0
21	0	0
20	0	0
10	$1-\epsilon$	0
01	0	1
200	$1-\epsilon$	0
100	0	1
000	0	1

Images taken from [A. Duarte, R. Fraiman, A. Galves, G. Ost and C.D. Vargas, arXiv:1602.00579]

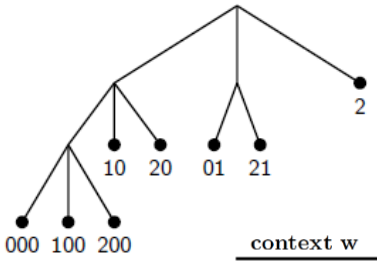
Retrieving a context tree form EEG data

Random Source



$\cdots 2\,1\,0\,1\,2\,1\,0\,1\,2\,1\,0\,1\,2\cdots$

$\cdots 2\,1\,0\,0\,2\,1\,0\,1\,2\,0\,0\,0\,2\cdots$



context w	p(1|w)	p(2|w)
2	$1-\epsilon$	0
21	0	0
20	0	0
10	$1-\epsilon$	0
01	0	1
200	$1-\epsilon$	0
100	0	1
000	0	1

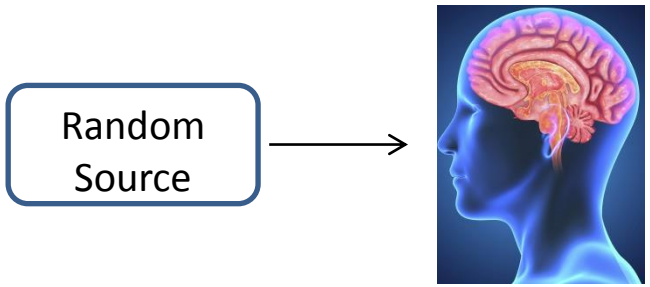
- $(X_1, Y_1), \ldots, (X_n, Y_n)$
- New statistical model selection procedure for FD
- The brain effectively identifies the context tree characterizing the source.

Images taken from [A. Duarte, R. Fraiman, A. Galves, G. Ost and C.D. Vargas, arXiv:1602.00579]

Can such finding be corroborated in behavioral responses, specifically in execution of movements?

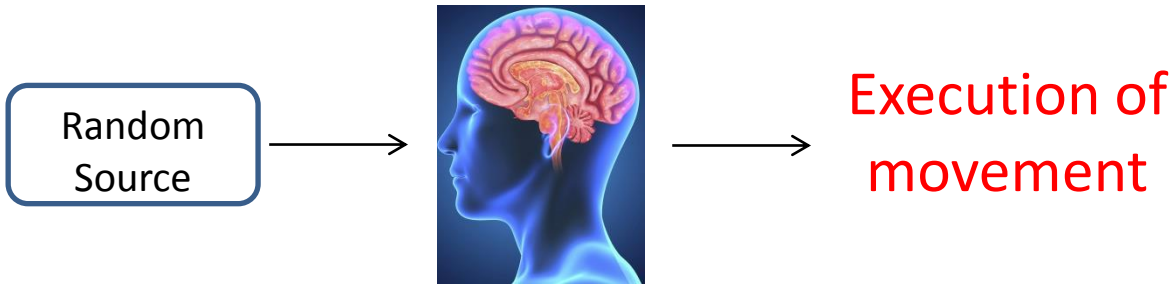
Images taken from
https://www.lvhn.org/for_referring_physicians/better_medicine/rehabilitation/neuromuscular_rehabilitation_technology_helps_patients_increase_motor_control; http://neuromat.numec.prp.usp.br/pt-br/newsletter-31; https://lehacker.com/brain-facts-revealing/

Can such finding be corroborated in behavioral responses, specifically in execution of movements?



Images taken from
https://www.lvhn.org/for_referring_physicians/better_medicine/rehabilitation/neuromuscular_rehabilitation_technology_helps_patients_increase_motor_control; http://neuromat.numec.prp.usp.br/pt-br/newsletter-31; https://lehacker.com/brain-facts-revealing/

Can such finding be corroborated in behavioral responses, specifically in execution of movements?



Images taken from
https://www.lvhn.org/for_referring_physicians/better_medicine/rehabilitation/neuromuscular_rehabilitation_technology_helps_patients_increase_motor_control; http://neuromat.numec.prp.usp.br/pt-br/newsletter-31; https://lehacker.com/brain-facts-revealing/

Can such finding be corroborated in behavioral responses, specifically in execution of movements?

Random Source →  → Execution of movement

Rehabilitation of patients

Goalkeeper Game

Images taken from
https://www.lvhn.org/for_referring_physicians/better_medicine/rehabilitation/neuromuscular_rehabilitation_technology_helps_patients_increase_motor_control; http://neuromat.numec.prp.usp.br/pt-br/newsletter-31; https://lehacker.com/brain-facts-revealing/

Stochastic processes driven by structured Markov chain

- Finite alphabet A
- Sequence of stimuli $X_n \in A$ generated according to specific regular statistical pattern

Stochastic processes driven by structured Markov chain

- Finite alphabet A
- Sequence of stimuli $X_n \in A$ generated according to specific regular statistical pattern
- Structured Markov chain $(X_n)_{n\in\mathbb{Z}}$:

Stochastic processes driven by structured Markov chain

- Finite alphabet A
- Sequence of stimuli $X_n \in A$ generated according to specific regular statistical pattern
- Structured Markov chain $(X_n)_{n\in\mathbb{Z}}$:
  - for some l,
  - any $m \geq l$, $(l, m \in \mathbb{Z})$
  - and any finite string $x_{n-m}^{n-1} = (x_{n-m}, \ldots, x_{n-1}) \in A^m$,

$$\begin{aligned} P\left(X_n = a | X_{n-m}^{n-1} = x_{n-m}^{n-1}\right) &= P\left(X_n = a | X_{n-l}^{n-1} = x_{n-l}^{n-1}\right) \\ &= P\left(X_n = a | c\left(x_{n-l}^{n-1}\right)\right) \text{ for all } a \in A. \end{aligned}$$

  - $c : A^l \to \mathcal{C} = \{c_1, \ldots, c_m\}$: mapping assigning to each past string a corresponding class in a partition of the space of relevant pasts.

Stochastic processes driven by structured Markov chain

- Sequence of response $Y_n \in \mathcal{Y}$ to the stimuli

Stochastic processes driven by structured Markov chain

- Sequence of response $Y_n \in \mathcal{Y}$ to the stimuli
- Stochastic processes driven by structured Markov chain $(X_n, Y_n)_{n \in \mathbb{Z}}$

Stochastic processes driven by structured Markov chain

- Sequence of response $Y_n \in \mathcal{Y}$ to the stimuli
- Stochastic processes driven by structured Markov chain $(X_n, Y_n)_{n\in\mathbb{Z}}$
  - $(X_n)_{n\in\mathbb{Z}}$ is a structured Markov chain.
  - $Y_1, Y_2, \ldots$ are independent variables conditionally to the sequence $(X_n)_{n\in\mathbb{Z}}$

$$P\left(Y_n \in J | X_{n-l}^n = x_{n-l}^n\right) = P\left(Y_n \in J | c\left(x_{n-l}^n\right)\right)$$

for any measurable J

Stochastic processes driven by structured Markov chain

- Sequence of response $Y_n \in \mathcal{Y}$ to the stimuli
- Stochastic processes driven by structured Markov chain $(X_n, Y_n)_{n\in\mathbb{Z}}$
  - $(X_n)_{n\in\mathbb{Z}}$ is a structured Markov chain.
  - $Y_1, Y_2, \ldots$ are independent variables conditionally to the sequence $(X_n)_{n\in\mathbb{Z}}$

$$P\left(Y_n \in J | X_{n-l}^n = x_{n-l}^n\right) = P\left(Y_n \in J | c\left(x_{n-l}^n\right)\right)$$

for any measurable J

Stochastic processes driven by context tree model is an outstanding example of SPDSMC

Related works: J.Garcia et al., *arXiv:1002.0729* (2010); V. Jääskinen et al., *Scand. J. Stat* (2014)

Prob1: Design of suitable experiments to evaluate evidence about learning statistical regularities

Prob1: Design of suitable experiments to evaluate evidence about learning statistical regularities

We need to carefully specify:

- An alphabet A
- Response variable Y_n

Prob1: Design of suitable experiments to evaluate evidence about learning statistical regularities

We need to carefully specify:

- An alphabet A
- Response variable Y_n

Goalkeeper Game

$A = \{0, 1, 2\}$ 0: center; 1: right; 2: left

Prob1: Design of suitable experiments to evaluate evidence about learning statistical regularities

We need to carefully specify:

- An alphabet A
- Response variable Y_n

Goalkeeper Game

$A = \{0, 1, 2\}$ 0: center; 1: right; 2: left

D_n Curves of spatial position, Videos gathered by sensors or camera

Prob1: Design of suitable experiments to evaluate evidence about learning statistical regularities

We need to carefully specify:

- An alphabet A
- Response variable Y_n

Goalkeeper Game

$A = \{0, 1, 2\}$

0: center; 1: right; 2: left

D_n

Curves of spatial position, Videos gathered by sensors or camera

$Y_n \in A$

direction

Prob1: Design of suitable experiments to evaluate evidence about learning statistical regularities

We need to carefully specify:

- An alphabet A
- Response variable Y_n

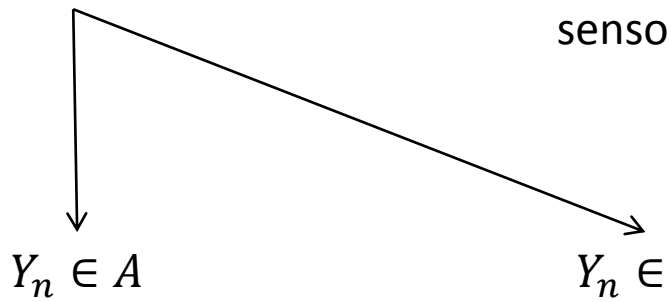
Goalkeeper Game

$A = \{0, 1, 2\}$

0: center; 1: right; 2: left

D_n

Curves of spatial position, Videos gathered by sensors or camera



$Y_n \in A$

direction

$Y_n \in \mathcal{Y}$

FD representing trajectory

Prob1: Design of suitable experiments to evaluate evidence about learning statistical regularities

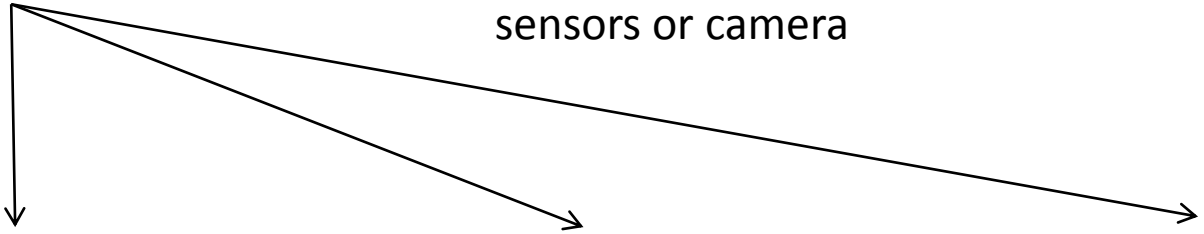
We need to carefully specify:

- An alphabet A
- Response variable Y_n

Goalkeeper Game

$A = \{0, 1, 2\}$ 0: center; 1: right; 2: left

D_n Curves of spatial position, Videos gathered by sensors or camera



$Y_n \in A$ $Y_n \in \mathcal{Y}$ $Y_n \in Y = \{0,1,2\} \times \{G, B\}$

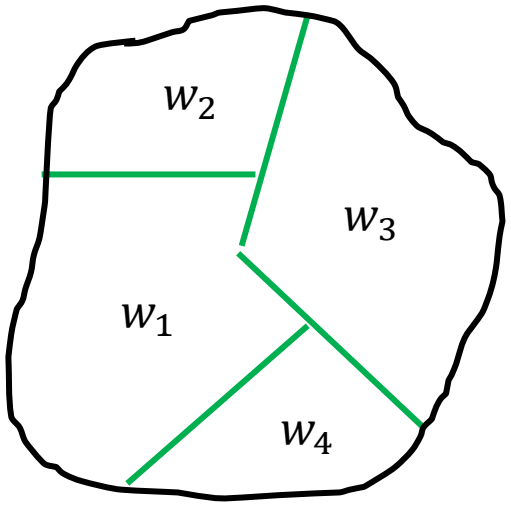
direction FD representing trajectory

- Kind of statistical pattern of the stimuli sequence $(X_n)_{n \in Z}$

  Context tree model compatible with (τ, p); $\quad p = \{p(\cdot \,|w) : w \in \tau\}$

- Kind of statistical pattern of the stimuli sequence $(X_n)_{n \in Z}$

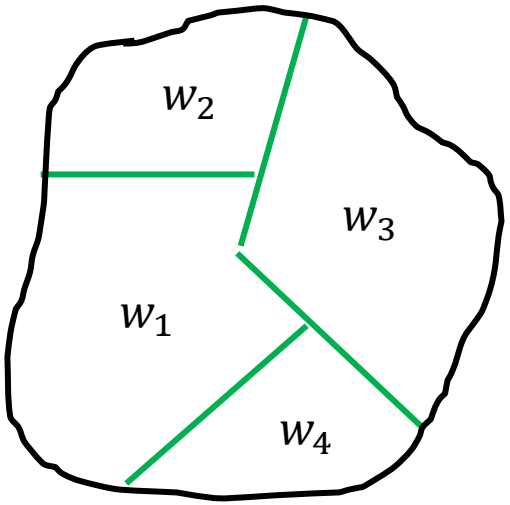
  Context tree model compatible with (τ, p); $\quad p = \{p(\cdot \,|w) : w \in \tau\}$



P_τ

- Kind of statistical pattern of the stimuli sequence $(X_n)_{n\in Z}$

Context tree model compatible with (τ, p); $\quad p = \{p(\cdot\,|w) : w \in \tau\}$



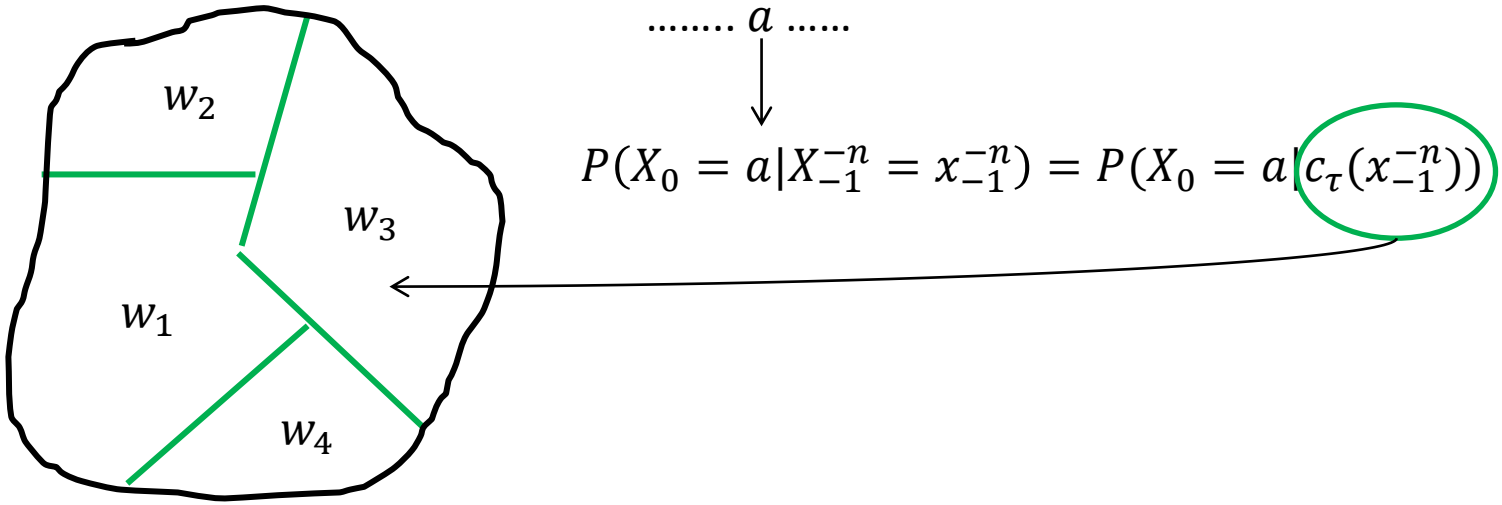
P_τ

$\ldots\ldots a \ldots\ldots$

$$P(X_0 = a | X_{-1}^{-n} = x_{-1}^{-n}) = P(X_0 = a | c_\tau(x_{-1}^{-n}))$$

- Kind of statistical pattern of the stimuli sequence $(X_n)_{n\in Z}$

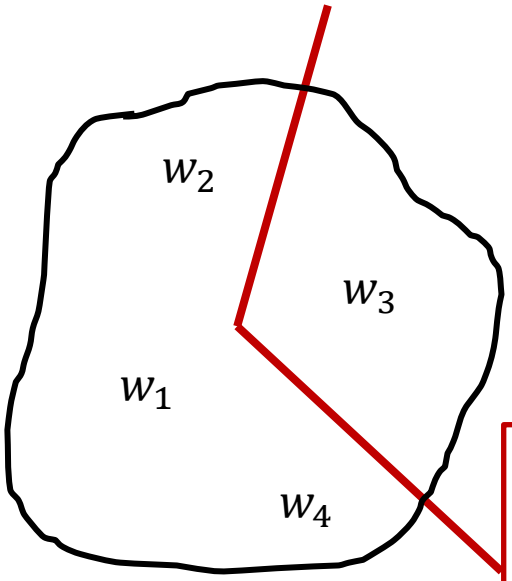
Context tree model compatible with (τ, p); $\quad p = \{p(\cdot\,|w) : w \in \tau\}$



P_τ

- Kind of statistical pattern of the stimuli sequence $(X_n)_{n \in Z}$

Context tree model compatible with (τ, p); $\quad p = \{p(\cdot\,|w) : w \in \tau\}$



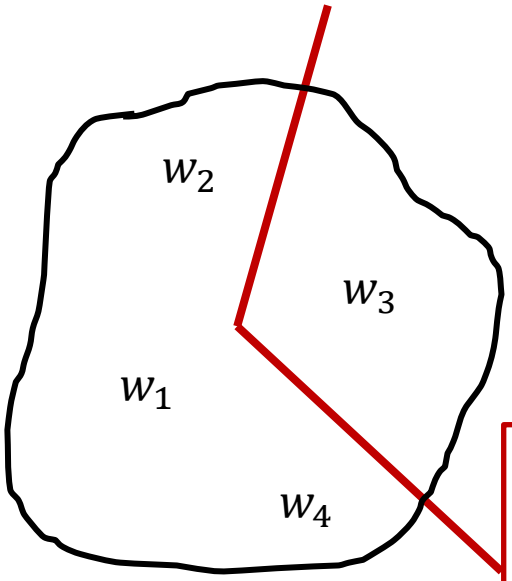
…… a ……

P_c : $s_1, s_2 \in A^*$ belong to the same element of P_C if

$$P(a|s_1) = P(a|s_2) \quad \forall a \in A$$

- Kind of statistical pattern of the stimuli sequence $(X_n)_{n \in Z}$

Context tree model compatible with (τ, p); $\quad p = \{p(\cdot\,|w) : w \in \tau\}$



…… a ……

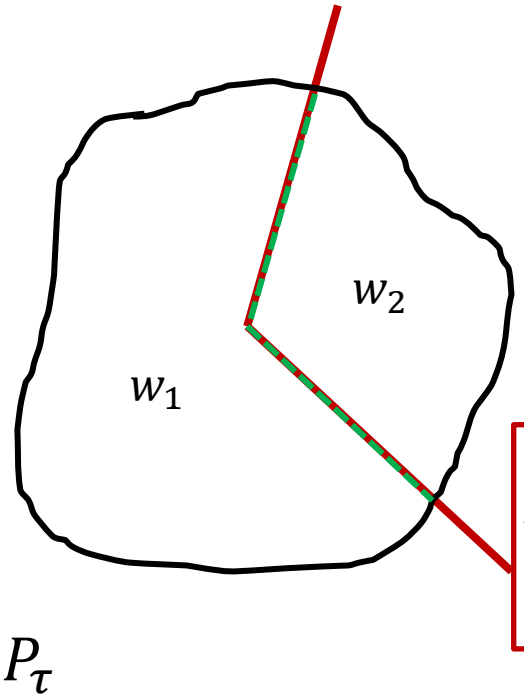
P_C : $s_1, s_2 \in A^*$ belong to the same element of P_C if

$$P(a|s_1) = P(a|s_2) \quad \forall a \in A$$

The subject might adopt a form of representation of the given Markov chain that involves a certain partition into classes that does not involves a context tree

- Kind of statistical pattern of the stimuli sequence $(X_n)_{n\in Z}$

Context tree model compatible with (τ, p); $\quad p = \{p(\cdot\,|w) : w \in \tau\}$



$\ldots\ldots a \ldots\ldots$

$$P(X_0 = a|X_{-1}^{-n} = x_{-1}^{-n}) = P(X_0 = a|c_\tau(x_{-1}^{-n}))$$

P_c : $s_1, s_2 \in A^*$ belong to the same element of P_C if

$$P(a|s_1) = P(a|s_2) \quad \forall a \in A$$

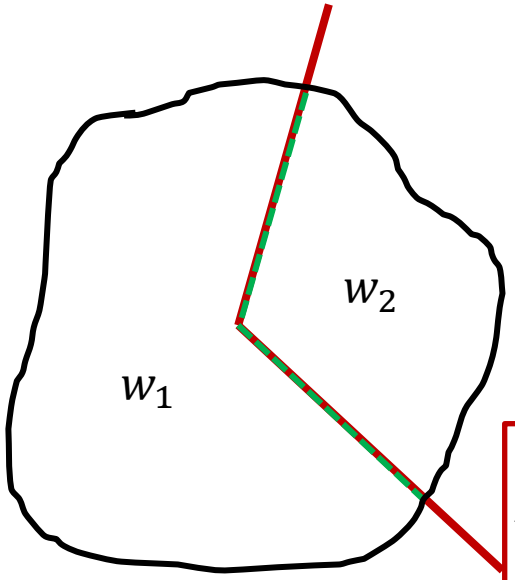
The subject might adopt a form of representation of the given Markov chain that involves a certain partition into classes that does not involves a context tree

- The transition probabilities measures $p = \{p(\cdot\,|w) : w \in \tau\}$ of the context tree model be different

$\Longrightarrow P_\tau = P_C$

- Kind of statistical pattern of the stimuli sequence $(X_n)_{n\in Z}$

Context tree model compatible with (τ, p); $\quad p = \{p(\cdot\,|w) : w \in \tau\}$



P_τ

$\ldots\ldots a \ldots\ldots$

$$P(X_0 = a | X_{-1}^{-n} = x_{-1}^{-n}) = P(X_0 = a | c_\tau(x_{-1}^{-n}))$$

P_c : $s_1, s_2 \in A^*$ belong to the same element of P_C if

$$P(a|s_1) = P(a|s_2) \quad \forall a \in A$$

The subject might adopt a form of representation of the given Markov chain that involves a certain partition into classes that does not involves a context tree

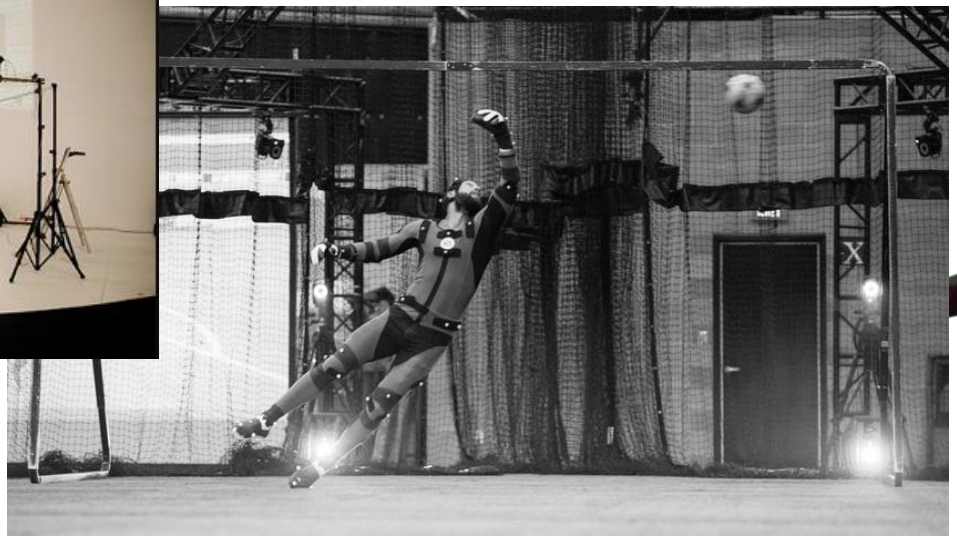
- The transition probabilities measures $p = \{p(\cdot\,|w) : w \in \tau\}$ of the context tree model be different $\quad \Longrightarrow P_\tau = P_C$

- How to quantify and control the complexity of the statistical pattern?

Prob2: Mapping from raw data to relevant response features Y_n

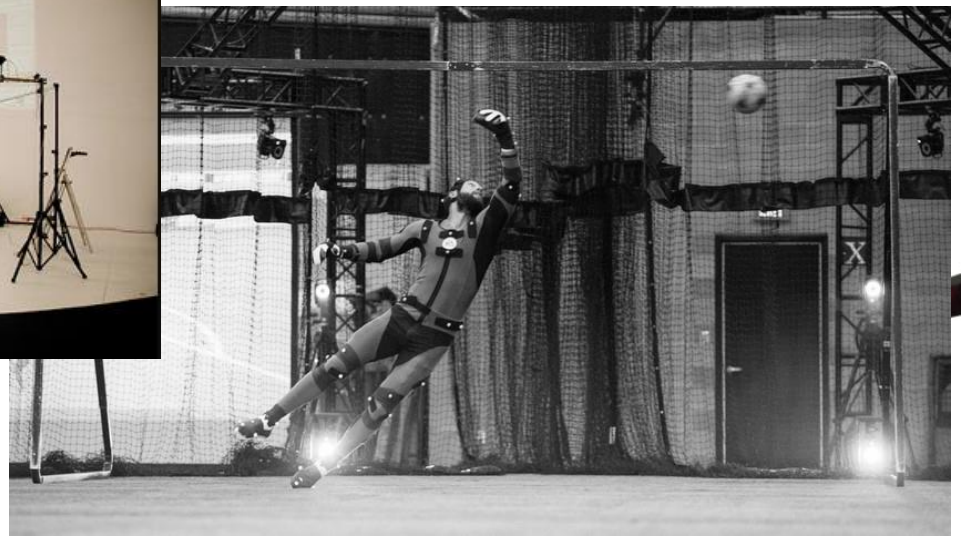
Prob2: Mapping from raw data to relevant response features Y_n

Motion Capture System: row data D_n



Prob2: Mapping from raw data to relevant response features Y_n

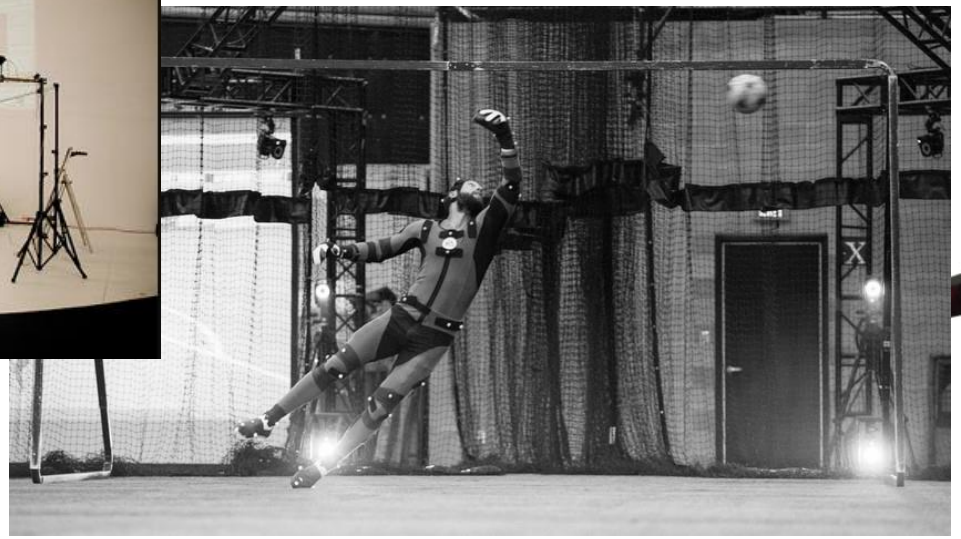
Motion Capture System: row data D_n



How to extract from such raw data relevant information Y_n concerning postures and gestures that characterize the response of the subject?

Prob2: Mapping from raw data to relevant response features Y_n

Motion Capture System: row data D_n



How to extract from such raw data relevant information Y_n concerning postures and gestures that characterize the response of the subject?

- Pattern recognition techniques and machine learning algorithms
  - Computer vision techniques
  - Gesture representation and recognition $\longrightarrow$ Y_n
  - Tracking algorithms

Prob3: Making inference about $(X_n, Y_n)_n$

For Y_n functional data:

A consistent model selection procedure

[A. Duarte, R. Fraiman, A. Galves, G. Ost and C.D. Vargas, arXiv:1602.00579]

Prob3: Making inference about $(X_n, Y_n)_n$

For Y_n functional data:

A consistent model selection procedure

[A. Duarte, R. Fraiman, A. Galves, G. Ost and C.D. Vargas, arXiv:1602.00579]

Two-sample hypothesis testing step

Prob3: Making inference about $(X_n, Y_n)_n$

For Y_n functional data:

A consistent model selection procedure
[A. Duarte, R. Fraiman, A. Galves, G. Ost and C.D. Vargas, arXiv:1602.00579]

Two-sample hypothesis testing step

Random projection method
[Cuestas-Albertos, Fraiman and Ransford (2006)]

Prob3: Making inference about $(X_n, Y_n)_n$

For Y_n functional data:

A consistent model selection procedure
[A. Duarte, R. Fraiman, A. Galves, G. Ost and C.D. Vargas, arXiv:1602.00579]

Two-sample hypothesis testing step

Random projection method
[Cuestas-Albertos, Fraiman and Ransford (2006)]

Two-sample distribution-free test for FD
[Pomann et al., *Appl. Statist.* (2016)]

Prob3: Making inference about $(X_n, Y_n)_n$

For Y_n functional data:

A consistent model selection procedure
[A. Duarte, R. Fraiman, A. Galves, G. Ost and C.D. Vargas, arXiv:1602.00579]

Two-sample hypothesis testing step

Random projection method
[Cuestas-Albertos, Fraiman and Ransford (2006)]

Two-sample distribution-free test for FD
[Pomann et al., *Appl. Statist.* (2016)]

Do not give information on specific aspects of the distributions causing detected differences.

Prob3: Making inference about $(X_n, Y_n)_n$

For Y_n functional data:

A consistent model selection procedure
[A. Duarte, R. Fraiman, A. Galves, G. Ost and C.D. Vargas, arXiv:1602.00579]

Two-sample hypothesis testing step

Random projection method
[Cuestas-Albertos, Fraiman and Ransford (2006)]

Two-sample distribution-free test for FD
[Pomann et al., *Appl. Statist.* (2016)]

Do not give information on specific aspects of the distributions causing detected differences.

Functional test that focus on particular parameters:

mean functions
covariance operators

Functional data analysis of motion curves

Biomedical questions of interest:

- Detection of deviation from normality in movements
- Longitudinal study of patients under treatment

Functional data analysis of motion curves

Biomedical questions of interest:

- Detection of deviation from normality in movements
- Longitudinal study of patients under treatment

Functional data analysis of motion curves

Biomedical questions of interest:

- Detection of deviation from normality in movements
- Longitudinal study of patients under treatment

- Functional cluster analysis
- Dissimilarity measures between functional populations
- Functional shape analysis
- Functional time series methods

Thank you