Joint Modelling of Behavioural and Brain Imaging Data

Peter Tiňo, Yuan Shen, Zoe Kourtzi, Rui Wang, Nahed Alowadi

PT - School of Computer Science
University of Birmingham
UK

me in brief

Foundational research in AI, Mach Learning, Natural Comp:
- learning theory
- state space models and dynamical systems
- coevolutionary dynamics, population based approaches to optimisation

Methodology advancements:
- metric learning
- blending modelling with machine learning

Inter- and cross-disciplinary research (extract science from data):
- astronomy
- cognitive neuroscience
- biomedical sciences
- computational finance

Behavioral experiment

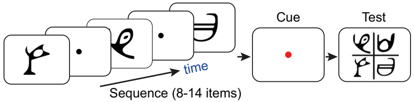
- 4 abstract symbols (retaining no semantic meaning/relationships)
- participants exposed to symbolic series over this alphabet generated by 0-, 1st- or 2nd- order Markov models.
- Markov models (actually VLMM) constructed in a hierarchical manner (to study possible transfer of information)
- after seeing a block of symbols there is a gap - participants are asked to predict what comes next (chose one of the 4 symbols presented in the screen in randomized positions).

Kourtzi, Wang, Shen, Tino, Welchman. The Journal of Neuroscience, 37(35), pp. 84128427, 2017.

Experimental setup

Figure 1

a

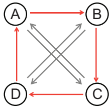


b

Level-0: Zero-order model

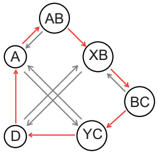
A	B	C	D
0.18	0.72	0.05	0.05

Level-1: First-order model



Level-1		Target			
		A	B	C	D
Context	A		0.8	0.2	
	B			0.8	0.2
	C	0.2			0.8
	D	0.8	0.2		

Level-2: Second-order model



Level-2		Target			
		A	B	C	D
Context	A		0.8	0.2	
	D	0.8	0.2		
	AB	0.2	0.8		
	XB			0.8	0.2
	BC		0.2	0.8	
	YC	0.2			0.8

Questions to investigate

- would like to investigate
  - transfer of knowledge across levels
  - transfer of learning capabilities across levels
  - higher level groupings of participants
  - strategies for coping with increasingly complex stimuli (environment) - probability matching/maximization
- BUT - sparse data difficult to capture - each participant saw a different series of symbols, albeit generated from the same source
- difficult to approach in a standard ML framework
- describe each participant (data item) by an associated model that captures his/her learning behaviour

Model that Tracks Participants' Learning

- Mixture of Markov models of order $k = 0, 1, 2$
- Predictive distribution for the t-th gap item:

$$p(s_t|s_{t-2}s_{t-1}) = w_0 \cdot p_0(s_t|\emptyset) + w_1 \cdot p_1(s_t|s_{t-1}) + w_2 \cdot p_2(s_t|s_{t-2}s_{t-1})$$

- Probability vector $\mathbf{w}$ is of central importance. Its time evolution $\mathbf{w}_t$ shows how a participant learned the correct set of contexts
- Bayesian approach to update $\mathbf{w}_t$.

Quantification Measures on the Models

- How well did a participant learn the underlying source?

$$\mathrm{KL}(\mathcal{M}^{true,k} || \mathcal{M}_t^{tracking,k}) = \sum_{c \in Context^k} p^k(c) \cdot \mathrm{KL}\left(p^k(\cdot|c) || p_t^k(\cdot|c) \right)$$

- Probability maximization? To what degree?

$$\mathrm{KL}(\mathcal{M}^{max,k} || \mathcal{M}_t^{tracking,k}) = \sum_{c \in Context^k} p^k(c) \cdot \mathrm{KL}\left(p^{max,k}(\cdot|c) || p_t^k(\cdot|c) \right)$$

where

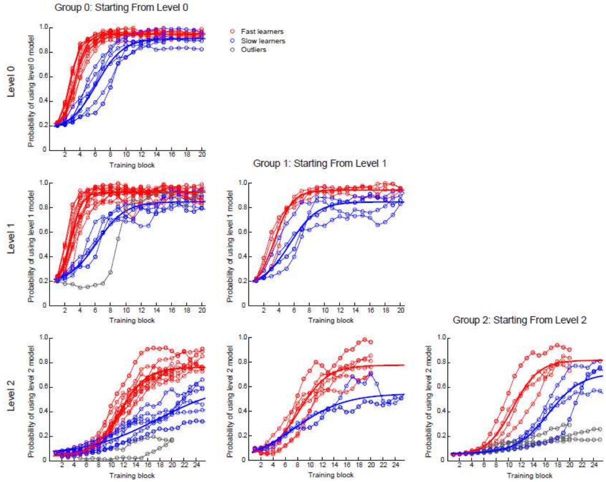
$$p^{max,k}(s|c) = \begin{cases} \approx 1 & s = s^{\mathbf{max}} \\ \approx 0 & \text{otherwise} \end{cases} \quad \text{with} \quad s^{\mathbf{max}} = \operatorname*{argmax}_s p^k(s|c).$$

- Probability maximization or matching?

$$\mathcal{D}_t^{strategy} = \mathrm{KL}(\mathcal{M}^{max,k} || \mathcal{M}_t^{tracking,k}) - \mathrm{KL}(\mathcal{M}^{true,k} || \mathcal{M}_t^{tracking,k})$$

Learning the Model Order

Figure 4



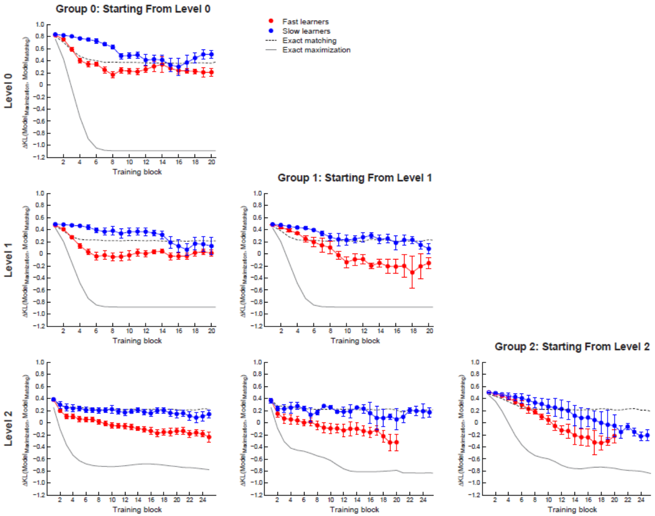
Key Question

Does the participants grouping based on learning the **model order** carry on to other aspects of behavioural studies?

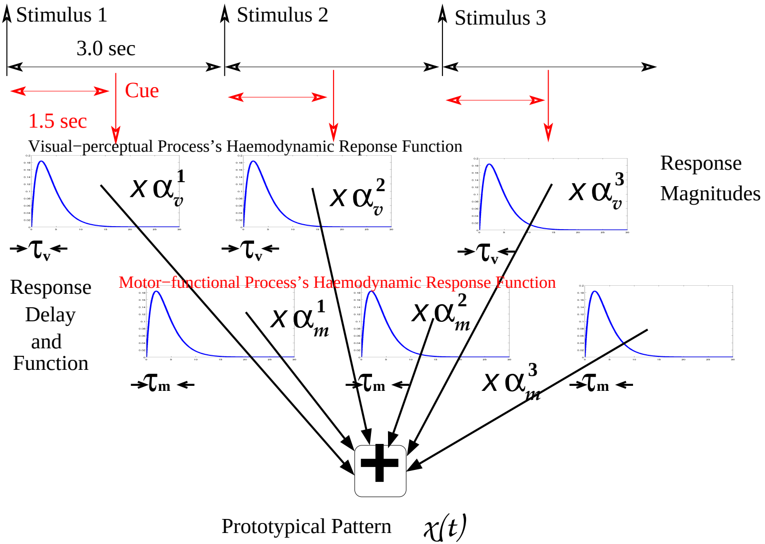
- transfer learning
- learning detailed source structure
- strategy choice (matching vs. maximization)

Strategy Choice

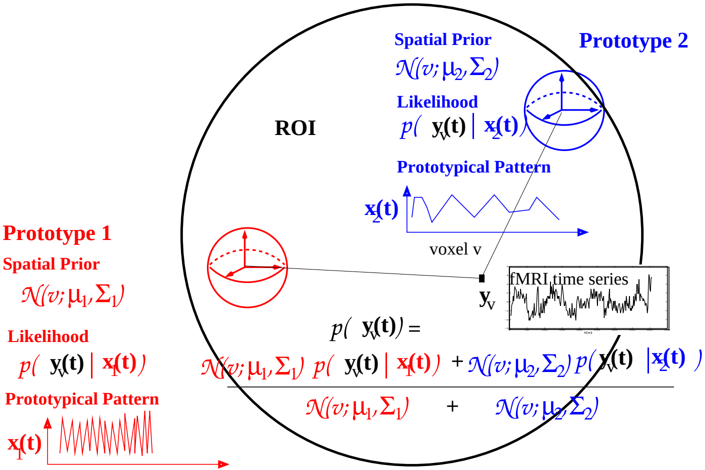
Figure 6



fMRI signals in ROIs - HPM modelling



Prototypes in space and time - ROI cortical activations



Shen, Mayhew, Kourtzi, Tino. Neuroimage, 84(1), pp. 657-671, 2014.

Simultaneous brain imaging and behavioral measurements

group level fMRI modelling

fast vs. slow learners

Bayesian formulation: **shared** tight **parameter priors in each group**

hierarchy of common/individualized within-group modelling

Shen, Alowadi, Wang, Kourtzi, Tino. Neural Comp, under review.

Regions of Interest

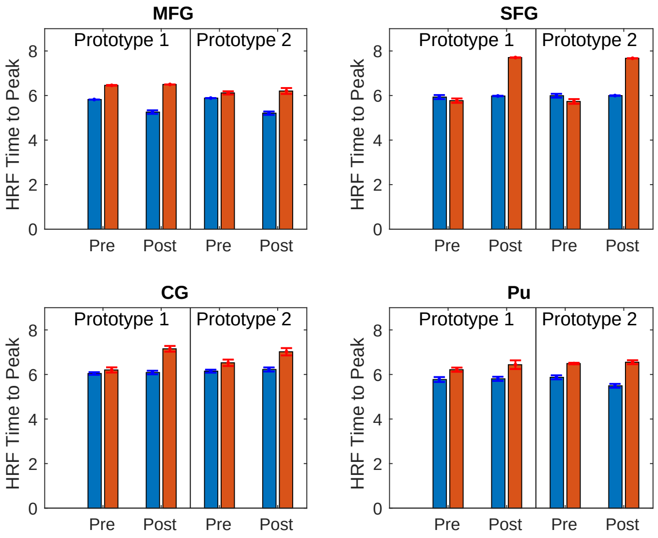
- **Frontal lobe**
  - **MFG** - middle frontal gyrus: attentional control.
  - **SFG** - superior frontal gyrus: coordination with sensory system.

  Both involved in learning, in particular, in sequential organization and self-monitoring of actions.

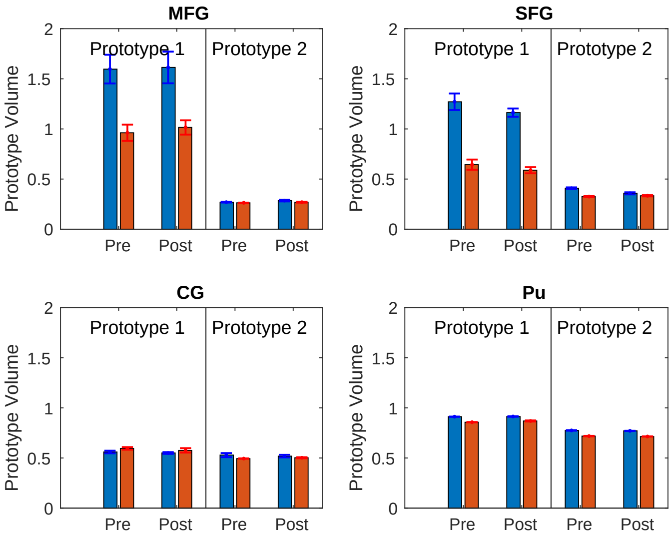
- **Limbic system**
  - **CG** - cingulate gyrus: integral part of the limbic system involved with emotion formation and processing, implicit learning and memory. Influential in linking motivational outcomes to behavior.

- **Basal ganglia**
  - **Pu** - putamen (lat. nutshell): Sub-cortical (dorsal stratum) associated with motor control, cognition, emotions and learning.

HRF - time-to-peak



Prototype volume



ROI homogeneity