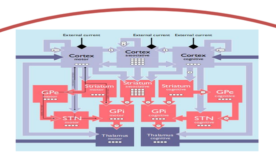
A neuro-computational model showing the effects of ventral striatum lesion on the computation of reward prediction error in VTA

Pramod S Kaushik, Frédéric Alexandre

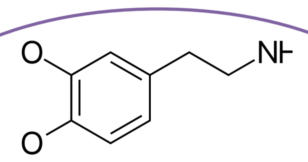
A neuro-computational model showing the effects of ventral striatum lesion on the computation of reward prediction error in VTA

Pramod KAUSHIK[1,2,3], Frédéric ALEXANDRE[1,2,3]

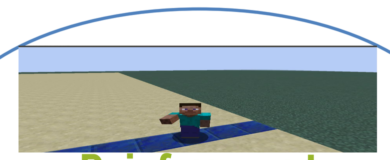
[1]INRIA Bordeaux Sud-Ouest, 200 Avenue de la Vieille Tour, 33405 Talence, France
[2]LaBRI, Université de Bordeaux, Bordeaux INP, CNRS, UMR 5800, Talence, France
[3]IMN, Université de Bordeaux, CNRS, UMR 5293, Bordeaux, France



Computational neuroscience



Dopamine



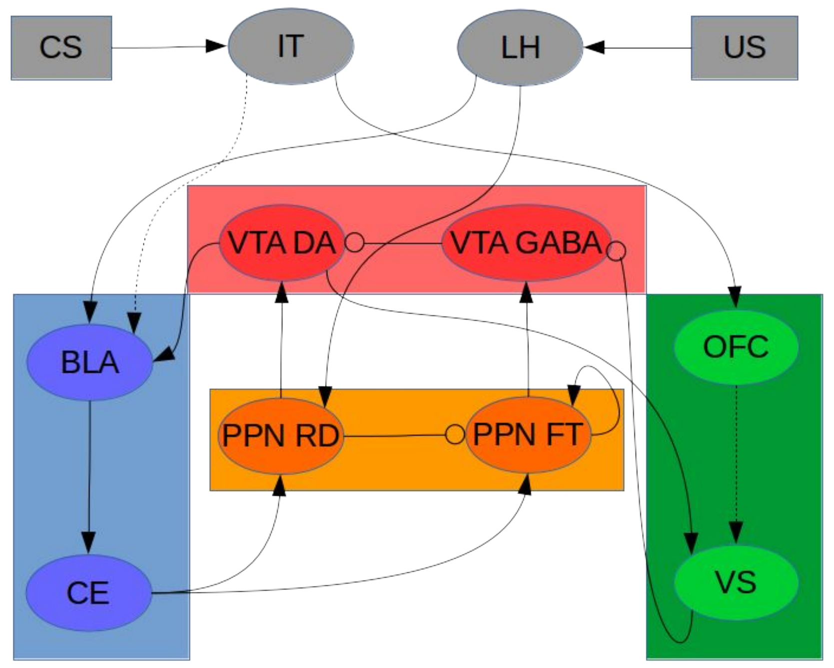
Reinformcent Learning

Abstract

- Modeling Pavlovian learning which is a fundamental learning mechanism in animals.
- Pairs a neutral conditioned stimulus (CS) with a rewarding unconditioned stimulus (US) and CS becomes rewarding stimulus after training
- Model focuses on the mechanism of RPE within pavlovian learning and the effects of Ventral Striatum (VS) lesions to illustrate a fundamental dissociation of magnitude and timing replicating experimental studies.

- Virtual lesions of VS to VTA GABA was made by disconnecting the link between them.
- Magnitude of reward is still conserved when lesions are made
- Timing information is lost indicating there are two dimensions to Pavlovian conditioning, namely timing and magnitude in a deviation from RL models

Model and Experimentation



Model Diagram illustrating structures involved in RPE Computation

Pointed arrows represent excitatory connections ,while rounded arrows represent inhibitory projections. Dashed lines represent learnable connections, while solid lines represent fixed connections in the model.

Computational Principles

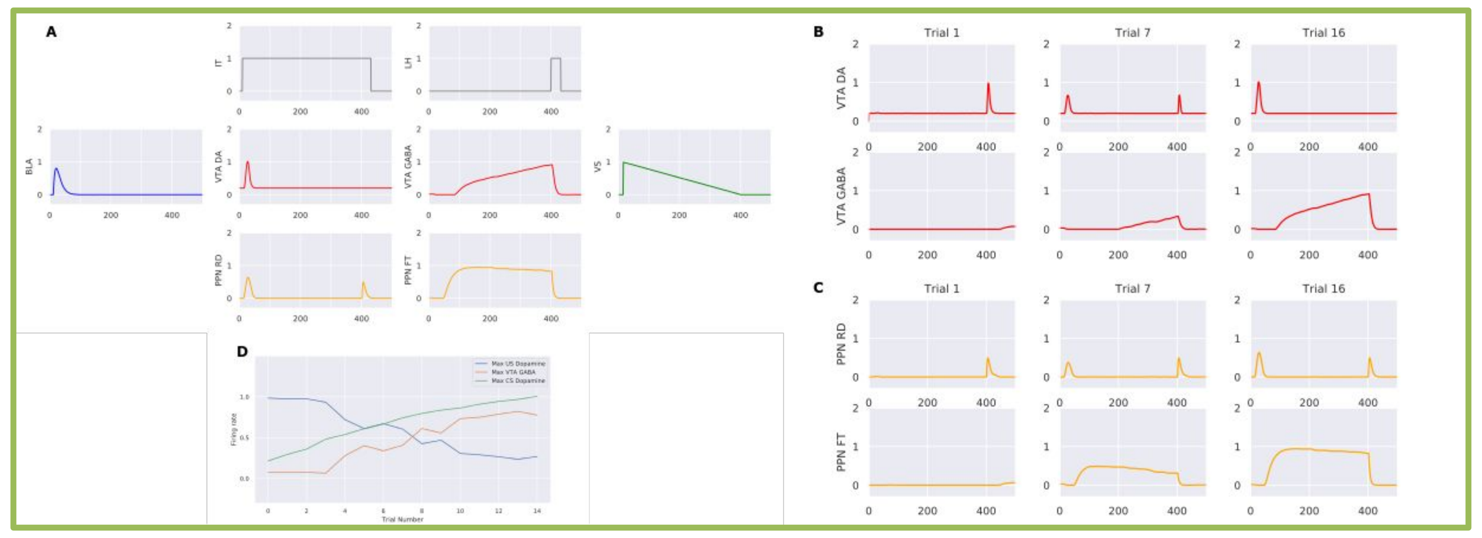
The proposed model is composed of computational units where each unit represents a population and computes the mean activity of the population.

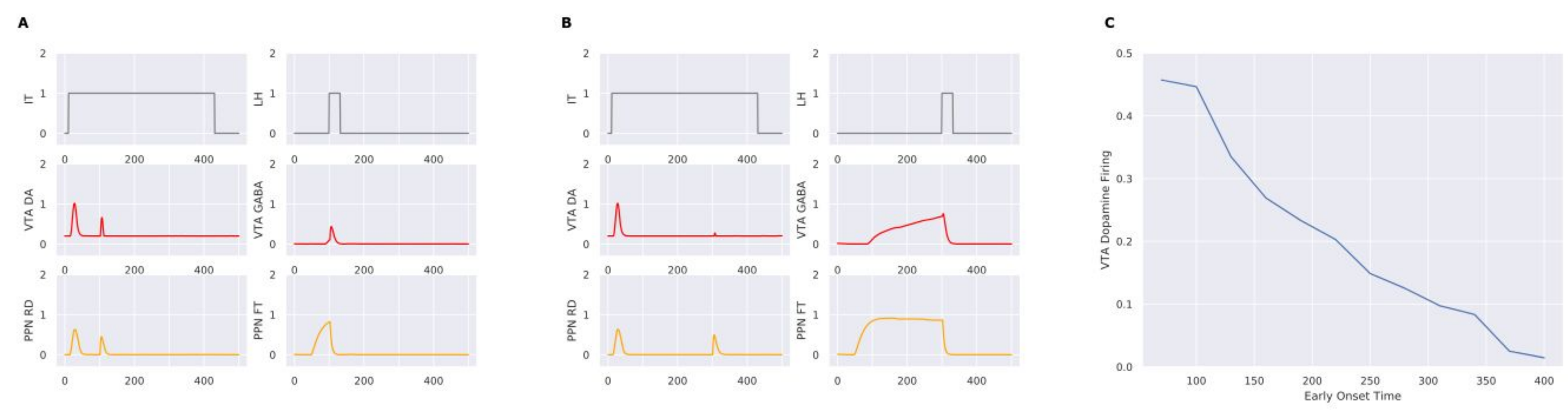
V(t) represents the membrane potential of the unit and the firing rate is a positive scalar of V(t) given by U(t).

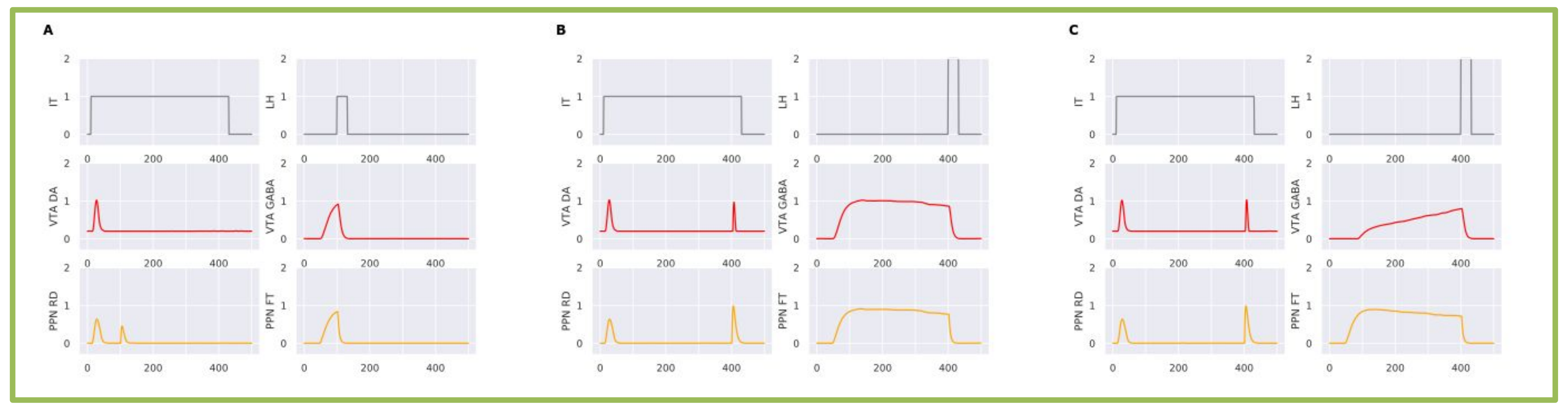
$$\tau.\frac{dV(t)}{dt} = (-V(t) + g_{exc}(t) - g_{inh}(t) + B + \eta(t))$$

$$U(t) = (V(t))^+$$

where τ is the time constant of the cell, B is the baseline firing rate and η(t) is the additive noise term chosen randomly at each time step from an uniform distribution between −0.01 and 0.01.







Model Terms	
VTA	*Ventral Tegmental Area*
VS	Ventral Striatum
LH	Lateral Hypothalamus
IT	Inferior temporal cortex
BLA	Basolateral Amygdala
CE	Central Amygdala
OFC	Orbitofrontal Cortex
PPN	Pedunculopontine nucleus

Features of the Model

- Portrays partial conditioning where VTA dopamine has acquired some CS firing and this expectation induces a partial expectation reducing US firing
- Not all early rewards have the same firing and sooner early rewards fire more than later early rewards in accordance with experiments (Fiorillo 2003)
- A new circuit with VTA GABA as a more biologically plausible expectation signal compared to VS (Keiflin 2015)
- VS Lesions do not affect magnitude encoding of the stimulus and only timing. (Takahashi 2017)

Predictions

- CE and PPN FT encode magnitude of expectation
- PPN through VTA GABA cancels dopamine
- A new circuit with VTA GABA as a more biologically plausible expectation signal compared to VS (Keiflin 2015)
- Early Reward cancellation of expectation happens within PPN
- Learning of Time before Learning of Magnitude

Conclusion and Acknowledgements

Conclusion

- Represents a model-free reinforcement learning system and learns the CS-US association in classical conditioning.
- Posits the brain could be solving the dimensions involved in classical conditioning separately in such a distributed manner.
- Such distributed processing could enable the same dimensions to be used to process other natural phenomena..

REFERENCES

Fiorillo, Christopher D., Philippe N. Tobler, and Wolfram Schultz. "Discrete coding of reward probability and uncertainty by dopamine neurons." Science 299.5614 (2003): 1898-1902.

Keiflin, Ronald, and Patricia H. Janak. "Dopamine prediction errors in reward learning and addiction: from theory to neural circuitry." Neuron 88.2 (2015): 247-263.

[Johnson et al., 2016] Johnson, M., Hofmann, K., Hutton, T., and Bignell, D. (2016). The malmo platform for artificial intelligence experimentation. In IJCAI, pages 4246–4247.

Takahashi, Yuji K., et al. "Temporal specificity of reward prediction errors signaled by putative dopamine neurons in rat VTA depends on ventral striatum." Neuron 91.1 (2016): 182-193.

Acknowledgements

We thank Bapi Raju in IIIT Hyderabad for constant guidance and feedback in this work.

NeuroFrance 2019, Marseille, France – 22-24 May, 2019