***Hippocampal state representations supporting imagination-based and habitual decision making systems: a large-scale computational model***

**David Samu, Universitat Pompeu Fabra**
***E-mail address:*** david.samu@upf.edu.

Several recent experimental findings have shed light on the functional anatomy of the mammalian brain's multiple decision making systems, particularly the deliberative internal simulation- (or imagination-) based system and the more automatic habitual system. At the same time, a wealth of theoretical and computational studies in the field of Reinforcement Learning have provided understanding about the relative merits and practical pitfalls of each approach, known as model-based and model-free decision making, respectively. Despite these advances, however, the interplay between these two systems is little known, and in particular, the inevitable competition and trade-off between them has been poorly studied. In this computational study, we model these decision making systems following our current knowledge of their large-scale functional anatomy and investigate the above questions in the context of a navigation task. To successfully solve the task, the simulated animal first has to learn to recognize its current state by integrating information from multiple noisy sensory sources, visual stimuli and proprioceptive feedback, mimicking the function of hippocampal place cells. An accurate enough model of the navigation environment allows the agent to mentally simulate the outcome of samples of potential action sequences, according to the experimentally observed imagination-based planning behaviour of rats during such tasks. Concurrently, a model of the dorsomedial striatum learns habitual state-action responses based on the outcomes of actions in each recognized states, and with learning gradually overtakes control from the deliberative systems. Competition between the two systems is induced by introducing non-stationarities to the environment, and the interplay between the systems is modelled based on a process tracking the animal's internal prediction error. This process allow the animal to fall back from a quick but inflexible habitual to a slow but flexible deliberative mode whenever the animal experiences surprise relative to its own expectations.