

The importance of neural diversity in complex cognitive tasks

Mattia Rigotti¹, Daniel Ben Dayan Rubin¹, Xiao-Jing Wang³, Stefano Fusi^{1,2}

¹Institute of Neuroinformatics, UNI—ETH Zurich, ²Center for Theoretical Neuroscience, Columbia Univ. New York, ³ Department of Neurobiology and Kavli Institute for Neuroscience, Yale Univ. School of Medicine, New Haven

Our decisions are usually based on a variety of factors, not only sensory inputs and response options, but also behavioral rule in play, previous choices and their consequences. All these factors determine the context which plays a key role in our decision process. Here we propose a model for the internal representation of contexts as attractor states of a neural circuit. Every contextual attractor corresponds to a specific inner mental state, which is self-sustained and reflects a specific disposition to behavior (e.g. how to interpret a stimulus in a specific context and to react to it in order to get reward). Every externally or internally generated event induces a transition from one inner mental state to another. In principle such a dynamical system can reproduce arbitrarily complex behaviors, provided that there are enough inner mental states. Notice that in the most general case the transition to a particular context state will depend on both the previous state (i.e. the attractor which is currently activated) and on the triggering event. We show that unfortunately the fraction of transitions that can actually be implemented in a dynamical system decreases rapidly with the total number of desired transitions (see Figure). Such a limitation does not depend on the representation of the attractors or the external events. The solution that we propose is to harness the huge diversity observed in the neural responses recorded in vivo, which most likely reflects the heterogeneity of the neural cells and their connections. If we start from the neurons which have a particular selectivity (e.g. to a stimulus, or to an intended motor response), other neurons with random connections with them would exhibit mixed selectivity. We prove that if there are enough of these mixed selectivity neurons, then any scheme of attractors and transitions can be implemented in a dynamical system. The recipe that we propose to build such a system is valid for any arbitrary neural representation of the inner mental states (attractors) and the events which trigger the transitions. We then illustrate the theory by building an attractor neural network of Hopfield like neurons which reproduces the animal behavior and several features of the neural activity recorded in prefrontal cortex in an experiment in which monkeys make decisions about motor response according to one of the two possible rules, and the rule currently in effect switch without warning [1].

Acknowledgments Supported by the SNF grant PP00A-106556 and by NIH-2R01 MH58754 **Refs** [1] Mansouri FA, Matsumoto K, Tanaka K., J. Neurosci. 2006

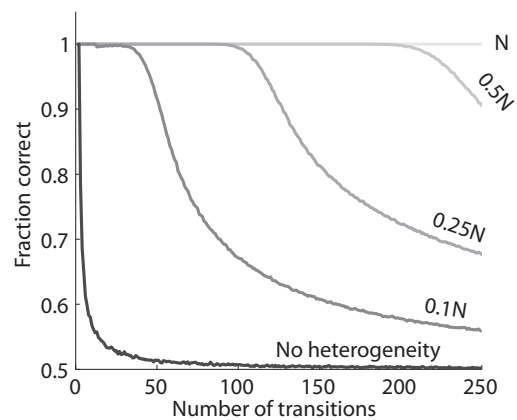


Figure 1: How the performance depends on heterogeneity: the performance is measured as the fraction of neural output units which satisfy the conditions for both the attractors and the transitions (both chosen randomly). Different curves correspond to different numbers of mixed selectivity neurons, MSNs, which measures the degree of heterogeneity. The number of MSNs is expressed a fraction (0, 0.1, 0.25, 0.5, 1) multiplied by the total number of units $N = 256$. For no heterogeneity (black curve), the performance drops quickly. Even a small number of MSNs can dramatically improve the performance. The number of implementable transitions grows linearly with the number of MSNs.