Computing with distributed distributional codes: convergent inference in brains and machines?

Maneesh Sahani[1]

1 Gatsby Computational Neuroscience Unit, UCL
E-mail: maneesh@gatsby.ucl.ac.uk

Abstract—A long-standing aspiration has been to discover a single theory that both underpins an understanding of computation in the nervous system, and forms the basis of effective machine learning and adaptive computation. However, despite the very substantial strides achieved in supervised learning over the past decade---and some hints of connections between the solutions found by such systems and neural represențions---there is still a sense that these machines work quite differently to brains. Part of the dissonance lies in the unrealism of backpropagation, and the undifferentiated structure of most "neural-network" architectures. But beyond that, the very problems being solved seem different. In particular, the brain's capacity for flexible inference: to parse and understand the components of an environment even if they are unfamiliar, or to reliably plan an action for the first time, appear to depend on learning a form of causal representation, most probably with little or no supervision. Such causal representation is also far more than the problem of density estimation that has dominated recent work in unsupervised machine learning, and which has similarly depended on backpropagation and unrealistic architecture.

After reviewing this situation I will outline recent work synthesizing a number of older ideas in both theoretical neuroscience and machine learning, which we hope will begin to lay the groundwork for a common theory of neural and machine inference.

Bio—Maneesh Sahani received both his BS degree (Physics) and PhD(Computation and Neural Systems) from the California Institute of Technology. He joined the UCL Gatsby Unit in 1999 with its first group of postdoctoral fellows, moved in 2002 to UCSF's Keck Center for Integrative Neuroscience for further postdoctoral work, before returning to the Gatsby Unit faculty in 2004. He is currently the Director of the Unit, and a Professor of Theoretical Neuroscience and Machine Learning.

His work explores the roles and uses of probabilistic inference in perception, neural processing and machine learning.