Statistics, Probability and Game Theory
IMS Lecture Notes – Monograph Series (1996) Volume 30

BAYESIAN UPDATINGS IN HOPFIELD-LIKE ASSOCIATE MEMORY MODELS

ISAAC MEILIJSON
Tel Aviv University

Abstract

This article summarizes and explains in statistical terminology two papers written jointly with Eytan Ruppin, presenting a Bayesian outlook on the performance of Hopfield-like attractor neural networks. Restricting attention to the evaluation of performance after two iterations rather than studying thermodynamical limits, we are able to extend the analysis to more general models than those usually considered: input patterns applied to small subsets of neurons, general connectivity architectures of the synaptic network and more efficient use of history. We show that the optimal signal that a Bayesian neuron should emit has a *slanted sigmoidal* shape as a function of its current field value (or posterior odds), and provide an intuitive account of activation functions with such non-monotone shapes.

1. Dedication. I studied in Berkeley at the Department of Statistics from 1967 to 1969. During this period I was a teaching and research assistant, mostly working under David Blackwell and Lester Dubins, to whom I equally owe the light that illuminates most of what I study. I had the privilege of being Blackwell's Ph.D. student, writing my thesis "A bargaining problem" under his guidance. I was also fortunate to be one of his teaching assistants in the course Stat 2, during the gestation period and writing of his marvelous textbook "Basic Statistics", the lucid, clear, elementary introduction to Bayesian thought. From Blackwell I learned stochastic modeling and Dynamic Programming, and got his Bayesian attitude into my bloodstream.

The following paper is warmly and gratefully dedicated to David Blackwell. It is a summary of an attempt with Eytan Ruppin to replace some ad-hoc dynamics transferred from the Statistical Mechanics language of spin models of Neural Networks, to a more Blackwellian approach: neurons have prior and posterior beliefs about their dichotomous hidden states, and these beliefs dictate their signals and decisions.

2. Introduction. In an *associative memory* model, there is a *storage phase* in which patterns are stored, and a *retrieval stage* in which distorted versions of the patterns are presented to the network, that is expected to recognize the correct input pattern by a sequence of "associations".