

Neural Networks & the Human Mind: New Mathematics Fits Humanistic Insight

Paul J. Werbos

Room 1151, National Science Foundation
Washington, D.C., USA 20550

Abstract The past two years have seen substantial progress in artificial neural networks (ANNs). New ANN control designs, combining reinforcement learning and generalized backpropagation [1,2], have demonstrated success on large-scale real-world problems which could not be solved by earlier designs, neural or nonneural. There now exists a ladder of designs, rising up from simple designs (of limited capability, but a good starting point), through to complex proven designs (which give more power and flexibility after they are mastered), up to new untested designs and ideas which should be able to replicate and explain human intelligence at the highest possible level.

After a brief review of neurocontrol, and an explanation of reinforcement learning, this paper asks what implications these designs have for our understanding of the human mind. It argues that this new mathematics is fully compatible with older deep insights into the human mind, due to humanistic thinkers East and West. One may therefore hope that this new view of the human mind may be of value in unifying important strands of human culture.

I. A REVIEW OF NEUROCONTROL

From an engineering point of view, the human brain is simply a computer, an information processing system. The function of any computer, as a whole system, is to compute its outputs. The outputs of the brain are control signals to muscles and glands. Therefore the brain as a whole system is a neurocontroller (a neural net control system) [3]. To understand the brain in functional, mathematical terms we should therefore focus on the subject of neurocontrol.

ANNs have been useful in four kinds of control task: (1) as subsystems of larger systems, where the controller itself is not an ANN; (2) in "cloning" applications, to copy what an expert does (unlike conventional expert systems which copy what an expert says); (3) in tracking applications, such as holding a plant to a fixed setpoint or making it follow a pre-specified reference trajectory; (4) in reinforcement learning or optimization over time. The first three of these four clearly have nothing to do with human intelligence.

Some biologists once argued that lower-level functions in the brain, such as the control of arm movement by the cerebellum, might be based on a simple tracking system; however, Kawato et al [4] have done impressive experiments proving that these parts of the brain actually optimize move-

ments over time. They do perform tracking, but only as part of an optimization task. A few authors have questioned these conclusions, but they still hold up quite well.

In brief, the reinforcement learning or optimization designs are the only designs of value in understanding the human mind. They are the only designs capable of meaningful planning or foresight. They are also the designs which have led to the most exciting real-world applications in recent years. Therefore, this paper will focus exclusively on them. To learn about other designs, stability, etc., see [1].

Within the field of optimization over time, two classes of design have proven useful in practice: (1) direct optimization, using generalized backpropagation to calculate derivatives of utility or performance or cost; (2) adaptive critic designs, which approximate dynamic programming.

Direct optimization based on the backpropagation of utility was first proposed in 1974 [2]. By 1988, there were four significant working examples [4], including two model robot controllers, one controller of a simulated truck-backer-upper, and a US Department of Energy official model of the natural gas industry [5]. By now, dozens of examples have appeared, including a Model-Predictive Control scheme now used to improve efficiency and reduce waste in profit-making chemical process plants, and an optimal tracking scheme of Narendra [1]. However, none of these designs are plausible as models of the brain. Some of the designs require calculations backwards through time; others require huge computational costs for large problems; and others simply cut off the key calculations which account for the long-term effect of present actions. These designs have great value in engineering (including an ability to reduce pollution while saving money), but they are not directly relevant to understanding the human mind; therefore, I will not discuss them here. Optimization methods derived from static function maximization are even less relevant here.

II. REINFORCEMENT LEARNING AND ADAPTIVE CRITICS IN GENERAL

Adaptive critic systems can perform "reinforcement learning." In reinforcement learning, an ANN system receives a vector of sensor inputs $\mathbf{X}(t)$. It outputs a vector of control signals or actions $\mathbf{u}(t)$. Then it receives a "reward" or "punishment" $U(t)$. In reinforcement learning, the system must somehow learn to output actions $\mathbf{u}(t)$ which maximize future rewards U , summed across all future time, from $U(t)$ to $U(\infty)$.

*The views herein are certainly not official views of NSF.

