Training a Limited-Interconnect, Synthetic Neural IC

Authors: M. Walker, S. Haghighi, A. Afghan, Larry Akers

Abstract: Hardware implementation of neuromorphic algorithms is hampered by high degrees of connectivity. Functionally equivalent feedforward networks may be formed by using limited fan-in nodes and additional layers. but this complicates procedures for determining weight magnitudes. No direct mapping of weights exists between fully and limited-interconnect nets. Low-level nonlinearities prevent the formation of internal representations of widely separated spatial features and the use of gradient descent methods to minimize output error is hampered by error magnitude dissipation. The judicious use of linear summations or collection units is proposed as a solution.

HARDWARE IMPLEMENTATIONS OF FEEDFORWARD,SYNTHETIC NEURAL SYSTEMSThe pursuit of hardware implementations of artificial neural network models is motivated by the need to develop systems which are capable of executing neuromorphic algorithms in real time. The most significant barrier is the high degree of connectivity required between the processing elements. Current interconnect technology does not support the direct implementation of large-scale arrays of this type. In particular. the high fan-in/fan-outs of biology impose connectivity requirements such that the electronic implementation of a highly interconnected biological neural networks of just a few thousand neurons would require a level of connectivity which exceeds the current or even projected interconnection density ofULSI systems (Akers et al. 1988).Highly layered. limited-interconnected architectures are however. especially well suited for VLSI implementations. In previous works. we analyzed the generalization and fault-tolerance characteristics of a limited-interconnect perceptron architecture applied in three simple mappings between binary input space and binary output space and proposed a CMOS architecture (Akers and Walker. 1988). This paper concentrates on developing an understanding of the limitations on layered neural network architectures imposed by hardware implementation and a proposed solution.778Walker, Haghighi, Afghan and AkersTRAINING CONSIDERATIONS FORLIMITED .. INTERCONNECT FEEDFORWARD NETWORKSThe symbolic layout of the limited fan-in network is shown in Fig. 1. Re-arranging of the individual input components is done to eliminate edge effects. Greater detail on the actual hardware architecture may be found in (Akers and Walker, 1988) As in linear filters, the total number of connections which fan-in to a given processing element determines the degrees of freedom available for forming a hypersurface which implements the desired node output function (Widrow and Stearns, 1985). When processing elements with

fixed, low fan-in are employed, the affects of reduced degrees of freedom must be considered in order to develop workable training methods which permit generalization of novel inputs. First. no direct or indirect relation exists between weight magnitudes obtained for a limited-interconnect, multilayered perceptron, and those obtained for the fully connected case. Networks of these types adapted with identical exemplar sets must therefore fonn completely different functions on the input space. Second, low-level nonlinearities prevent direct internal coding of widely separated spatial features in the input set. A related problem arises when hyperplane nonlinearities are used. Multiple hyperplanes required on a subset of input space are impossible when no two second level nodes address identical positions in the input space. Finally, adaptation methods like backpropagation which minimize output error with gradient descent are hindered since the magnitude of the error is dissipated as it back-propagates through large numbers of hidden layers. The appropriate placement of linear summation elements or collection units is a proposed solution.