Fermi National Accelerator Laboratory

FERMILAB-Conf-90/21

Status of HEP Neural NET Research in the U.S.A.*

Bruce Denby
Fermi National Accelerator Laboratory
P.O. Box 500
Batavia, Illinois 60510

and

Stephan L. Linn
Florida State University
Tallahassee, Florida 32306

January 1990

* Talk presented by B. Denby at the 1989 Conference on Computing in High Energy Physics, Oxford, England, April 10-14, 1989.

Operated by Universities Research Association Inc. under contract with the United States Department of Energy

STATUS OF HEP NEURAL NET RESEARCH IN THE USA †

Bruce Denby
Fermi National Accelerator Laboratory *
Batavia, IL 60510

Stephan L. Linn
Supercomputer Computations Research Institute
Florida State University, Tallahassee, FL 32306

ABSTRACT

Progress on tracking with recurrent neural networks is presented. Applications of feed forward networks to High Energy Physics are discussed. The situation regarding hardware implementations of neural networks is assessed.

1. Track Reconstruction

In previous work [1] we have discussed a method for reconstructing tracks using a Hopfield style recurrent neural network. Directed links between hit points were identified with neurons which interact with each other via an interconnection strength

$$T_{ij} = \frac{A cos^n \theta_{ij}}{l_i l_j}$$

where θ_{ij} is the angle between links i and j, and l_i, l_j are their lengths. It is interesting to note that an essentially identical connection matrix was derived independently [2] by Peterson.

Considerable effort was put into trying to optimise the performance by varying the form of the T_{ij} matrix. Following an argument of Zucker [3] we believe that one will probably never obtain perfect performance without somehow incorporating curvature information. Up until now our search of parameter space has been rather haphazard, as learning algorithms for recurrent networks such as that of Pineda [4] were unknown to us. Although we do not expect to achieve perfect performance, we intend to use such a learning rule to deduce an optimum connection matrix T_{ij} and compare it to the one we obtained heuristically.

The neural net performs a useful function even though it does not suppress all of the incorrect links. That this is so can be seen by comparing a Hough transform of the data before and after the application of the neural net algorithm, as is shown schematically in figure 1. The neural net, by using only local slope information has improved the signal to noise ratio for tracks over background by an large factor. The final track information could be efficiently obtained by presenting the neural net results to a track fitting program.

In our previous work, we suggested that the neural net algorithm could provide big speedups if implemented on vector or parallel machines. We have begun to look into this question [5]. Simulated events of the form $W \to jet\, jet$ were generated for a detector modelled on the D0 detector at Fermilab. Each charged particle could form as many as six generic 'hits'; one in the TRD, one in each of four cylindrical layers of tracking chambers (called the CDC; these 'hits' are actually segments but we did not make use of this information),

† Talk presented at the conference Computing in High Energy Physics, Oxford, U.K., 10-14 April, 1989.

* Fermilab is operated by Universities Research Association, Inc., under contract with the U.S. Department of Energy.

and one each in the electromagnetic and hadronic calorimeters outside the CDC. Links were defined between hits as before but this time the connection matrix used was

$$T_{ij} = f_{ij} cos^n \alpha_{ij} cos^n \beta_{ij}$$

where α_{ij} is the angle between i and j placed tail to tail, β_{ij} is the angle between radial vectors to the midpoints of i and j, $n = 100$, f_{ij} is set equal to $+1$ for i, j head to tail, -1 for i, j head to head or tail to tail, and 0 in all other cases.

Sample results are shown in figure 2. On a small sample of events, it was possible to achieve perfect reconstruction by deleting all neurons with activations less than .875 (maximum = 1). The routine was implemented both in both scalar and vector versions on a Cyber 205 vector supercomputer. The vectorized version ran a factor of 20 faster than the scalar version. It is interesting to note that the presence or absence of links from the tracking chamber to the calorimeter can be interpreted as a particle's 'pointing' or 'not-pointing' to a calorimetric energy deposit and thus may be of use in particle identification.

II. Feed-forward Networks

A number of interesting applications of feed-forward networks in High Energy Physics have appeared. In a feed-forward network, the neurons are arranged into an input layer which accepts the data, a number of 'hidden' layers, and an output layer which encodes the answer (figure 3). Connections are made only from lower layers to higher layers; there are no connections within layers and no feedback. Such networks can be used for a variety of pattern classification tasks. The necessary T_{ij} can be obtained with 'learning rules'. A popular one is called 'back-propagation'. In this procedure, a χ^2 is defined which is proportional to the 'correctness' of the network's performance:

$$\chi^2 = \sum_t \sum_i (V_{t,i} - \xi_{t,i})^2$$

where $V_{t,i}$ is the activation of neuron i in event t, and $\xi_{t,i}$ is the desired activation of neuron i in event t. One minimizes the χ^2 through gradient descent:

$$\Delta T_{ij} = -\epsilon \frac{\partial \chi^2}{\partial T_{ij}}.$$

Back propagation learning in feed forward networks has been used for e/γ discrimination in a layered calorimeter, and for γ/jet discrimination in a multicell calorimeter [6]. Planned future activities are identification of electrons in jets and track segment finding in SSC-style straw chambers [7]. There are a number of commercial neural net simulation packages available on the market, and the above studies have been carried on some of these [8]. Most run on IBM PC's, and source code is not available, which is rather inconvenient since one needs to run event generators on bigger machines and to modify the programs at will. Efforts are currently underway to concoct a simulator more amenable to the needs of High Energy Pysicists.

III. Hardware 'Progress'

A neural net tracking processor for a high resolution 4π tracking system would require of the order of 10^8 interconnections. Such high density can probably only be realized in VLSI or with optical methods. The status of large scale, truly parallel, hardware neural net implementations has not changed in the last year: one only hears about a few experimental devices. The reasons for this are not clear. It is probably true that in many applications the time and expense of VLSI is not warranted since the speed of neural networks is not very important for things like speech processing or handwritten character recognition: these can be run as simulations. There are certain problems where the speed is very important, such as in SDI type applications like missile tracking, but this work may be classified, and that could be why we are not hearing about it. It may turn out that HEP will have to be a pioneer in large scale neural network hardware.

The applications to calorimetry require far lower numbers of neurons (one per calorimeter cell) and connections ($n_{input} * n_{hidden} + n_{hidden} * n_{output}$). As such, these may be good candidates as a first application of hardware neural nets as a proof of technique and to gain experience.

For tracking, where component counts will be much higher, one wants to find an application which is small enough to be feasible with existing technology, but large enough to solve a real problem in an impressive way. We have not yet found such an application, though the search continues. Even for a small application, one will probably have to 'cheat' by doing some of the calculation with DSP's or microprocessors. This will slow down the operation of the nets but hopefully not enough to make them uninteresting. Another line of attack currently under study for tracking is to treat the position of a hit with respect to a wire (i.e., the drift time) as an analog quantity. This allows to reduce granularity of the system and thereby the total number of neurons [7].

Figure Captions

Figure 1. The top frame shows two straight tracks each formed of five hits (crosses). After neural net evolution, the correct links (solid lines), and a few incorrect ones, remain. In the Hough transform, we histogram the angles of (dashed) lines from each intercept bin to each hit (lower frame). In the top curve, before neural net evolution, there is much combinatorial background. After evolution, we histogram only those lines consistent with valid links; most of the background disappears.

Figure 2. Dotted lines show locations of generated tracks. 'Generic' hits are represented by crosses. Solid black lines are found links after convergence of the network. Some incorrect links are visible, but these can be removed by the cut mentioned in the text.

Figure 3. A typical three layer feed forward net. This figure was produced by the authors of reference 6 using the simulator of reference 8.

References

1) B. Denby,*Neural Networks and Cellular Automata in Experimental High Energy Physics*, *Computer Physics Communications* **49** (1988) p. 429-448.

B. Denby, *Neural Network and Cellular Automata Algorithms*, FSU-SCRI-88-141, Florida State University, Tallahassee, FL, June 1988.

2) C. Peterson, *Track Finding with Neural Networks*, LU TP 88-8, University of Lund, Sweden, submitted to *Nuclear Instruments and Methods*, April, 1988.

3) Pierre Parent and Steven W. Zucker, *Trace Inference, Curvature Consistency, and Curve Detection*, McGill University preprint CIM-86-3, June, 1985.

1) Pineda, F., *Generalization of Back Propagation to Recurrent Neural Networks*, *Physical Review Letters* **59** (1987) p. 2229.

5) Bruce Denby and Stephan L. Linn, *High Energy Particle Identification using a Neural Network Algorithm*, in preparation.

6) C. Barter et al.,*Neural Networks, D0, and the SSC*, presented at the Conference on Triggering and Data Acquistion for Experiments at the Supercollider, Toronto, Canada (January, 1989).

7) Bruce Denby et al., work in progress.

8) A typical one is the *Neuralworks Professional II^{TM}*, available from NeuralWare Inc., 103 Buckskin Court, Sewickley, PA 15143 U.S.A..

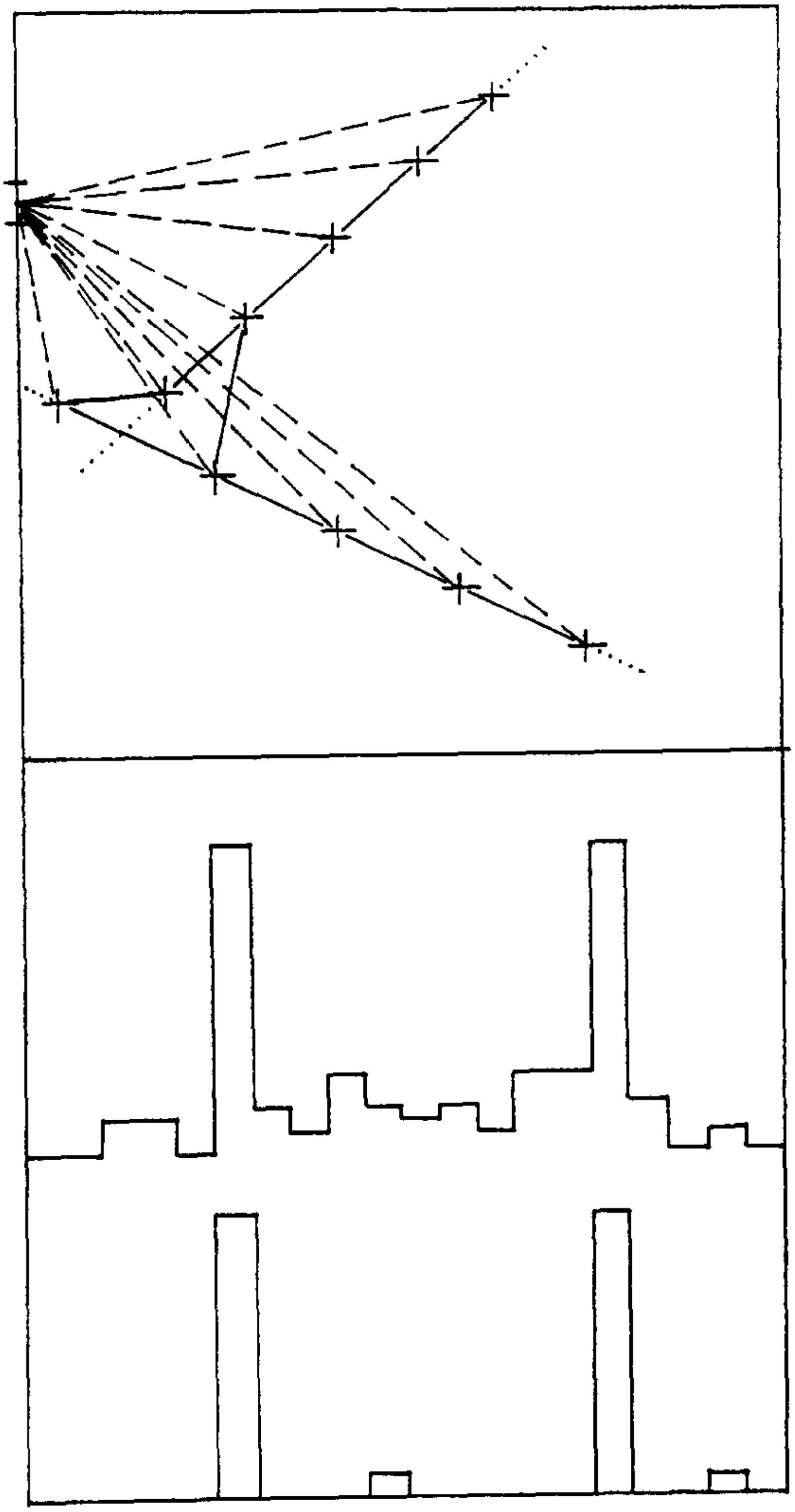
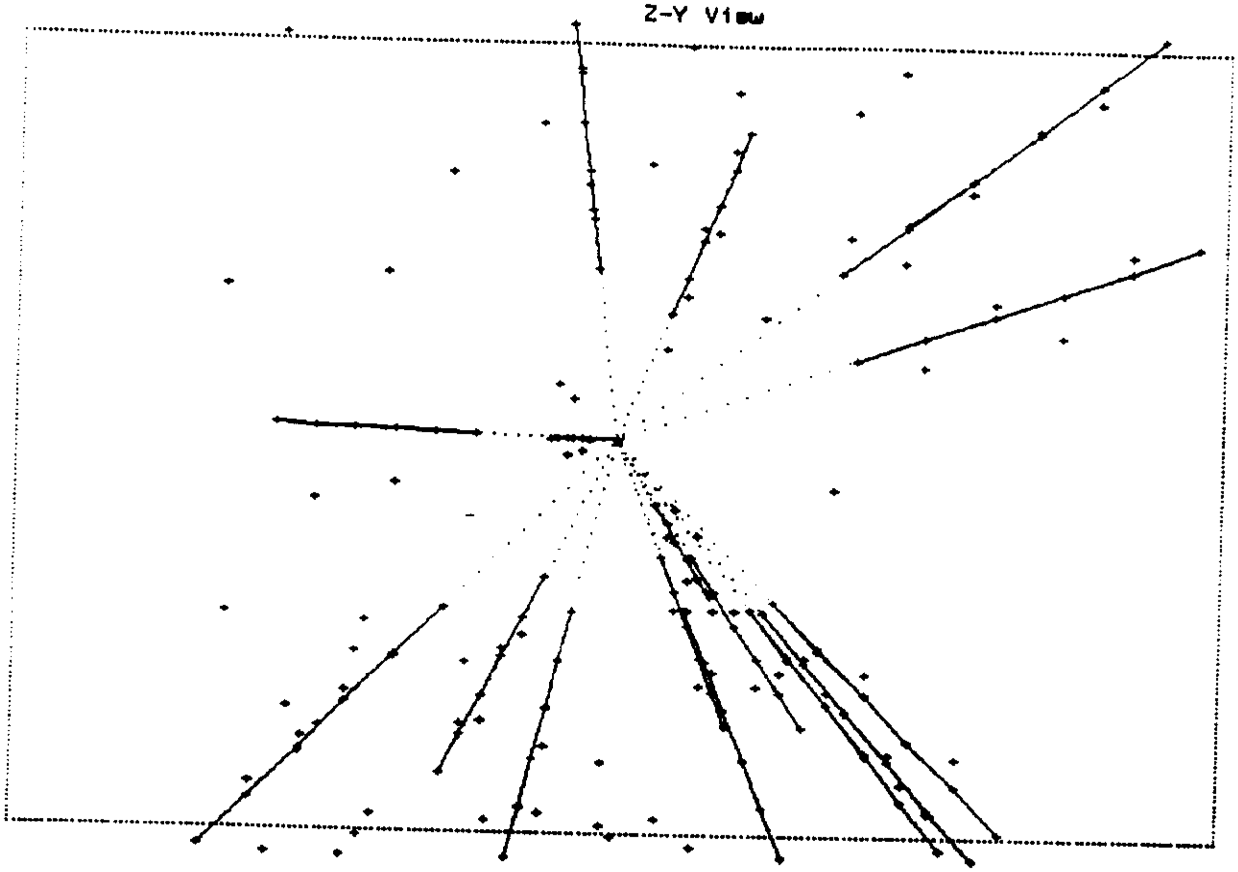


Figure 1



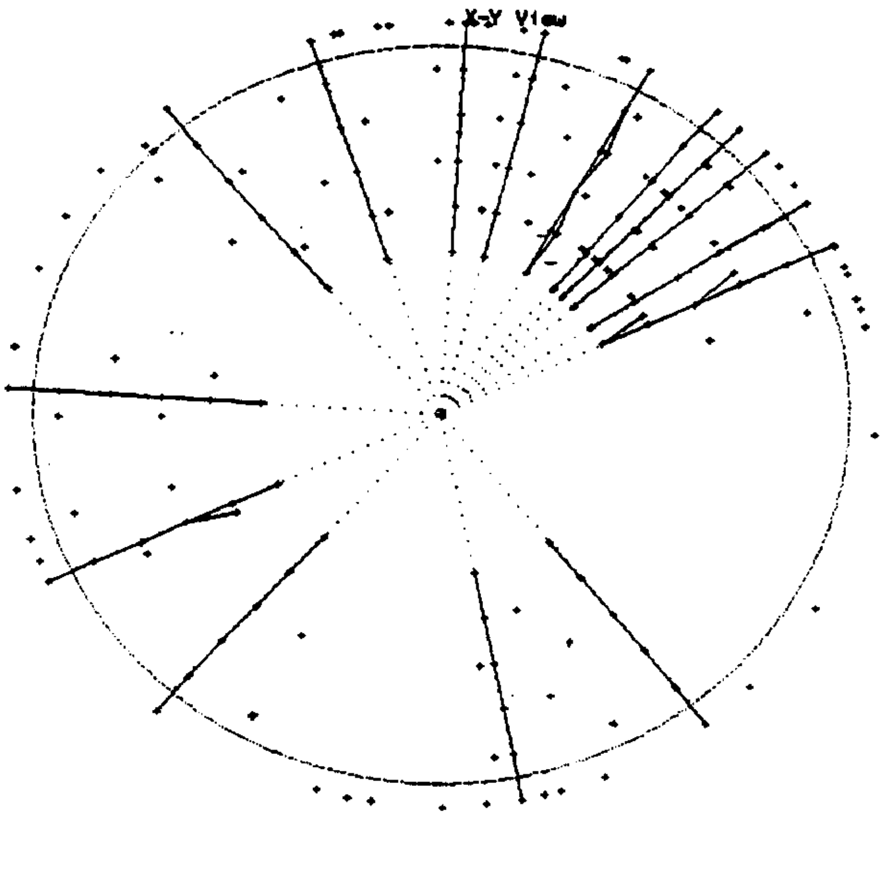


Figure 2

Instanet Standard Back-Propagation Network version 1.00 20-Jun-88

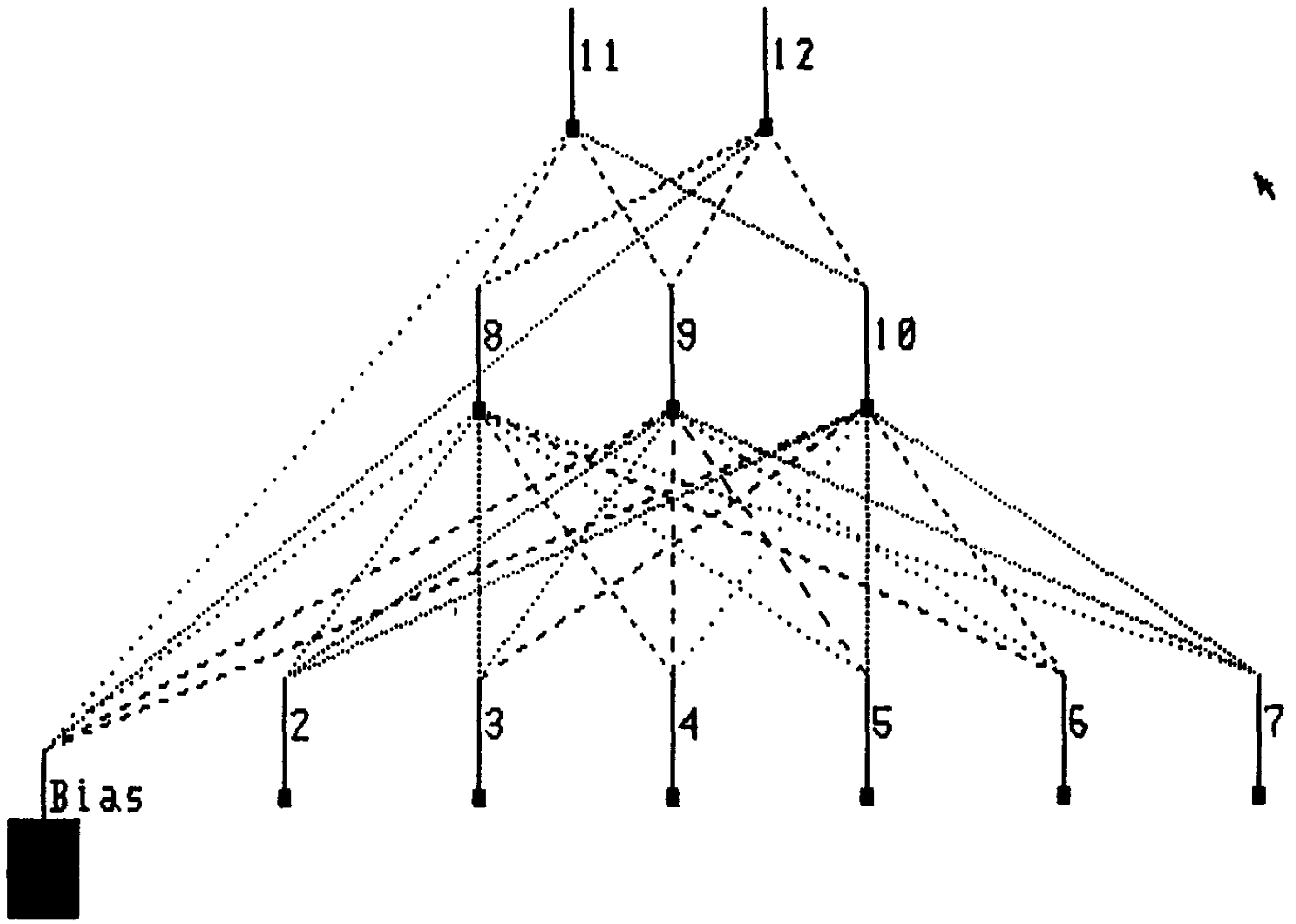


Figure 3