The H1 Neural Network Trigger Project

C. Kiesling[1]*, B. Denby[2], J. Fent[1], W. Fröchtenicht[1], P. Garda[2], B. Granado[2], G. Grindhammer[1], W. Haberer[1], L. Janauschek[1], T. Kobler[1], B. Koblitz[1], G. Nellen[1], J.-C. Prevotet[2], S. Schmidt[1], E. Tzamariudaki[1], S. Udluft[1]

[1] *Max-Planck Institut für Physik, München, Germany*
[2] *Laboratoire des Instruments et Systemes, Universite Pierre et Marie Curie, Paris VI, France*

Abstract. We present a short overview of neuromorphic hardware and some of the physics projects making use of such devices. As a concrete example we describe an innovative project within the H1-Experiment at the electron-proton collider HERA, instrumenting hardwired neural networks as pattern recognition machines to discriminate between wanted physics and uninteresting background at the trigger level. The decision time of the system is less than 20 microseconds, typical for a modern second level trigger. The neural trigger has been successfully running for the past four years and has turned out new physics results from H1 unobtainable so far with other triggering schemes. We describe the concepts and the technical realization of the neural network trigger system, present the most important physics results, and motivate an upgrade of the system for the future high luminosity running at HERA. The upgrade concentrates on "intelligent preprocessing" of the neural inputs which help to strongly improve the networks' discrimination power.

INTRODUCTION

It was here at Fermilab, in the early nineties, where the first applications of hardwired neural networks in high energy physics were initiated. The goal of these projects was to construct fast, intelligent triggers for HEP experiments. At that time, only a few neuromorphic chips were available, mostly in the form of analog devices. The early Fermilab applications all used such analog chips. Since then a number of digital instances have been developed, opening up the field for concrete neural trigger applications, which can be bit-precisely simulated offline for controlled determination and monitoring of trigger efficiencies. In these applications it has become clear that the digital approach is essential for further progress of the neural technology in high energy physics.

While many successful neural applications have emerged during the past years in the field of offline analysis (see, e.g., reports at this conference such as (1, 2)), the trigger area is still highly sensitive and new methods such as neural networks, although understood theoretically in great detail, should be prepared for quite some skepticism. Nevertheless, a number of successful projects have emerged and are turning out now new physics results, hard to obtain with the more "classical" methods of triggering. The main subject of this article is to show, giving the concrete example of the H1 neural network trigger, that the neural method in the trigger area is now firmly established, provides new physics information, and certainly has a splendid future in areas where fast and efficient pattern recognition algorithms are mandatory.

NEUROMORPHIC HARDWARE AND EARLY HEP APPLICATIONS

Neural networks have traditionally been studied in a simulation approach on standard serial computers, where the inherent parallelism of the neural method is not exploited. These more academic applications are usually not time-critical and can afford an execution speed at the "millisecond" scale (we speak here about the computations of a multi-layer feed-forward net in its "recall" step. The training of networks, which is done beforehand in the standard backpropagation supervised mode, needs many order of magnitude more execution time, depending on the amount of training data available and on the complexity of the network architecture). In the early 90's first neuromorphic hardware (massively parallel processor architectures) became available from industry so that high speed ("microsecond scale") realtime applications, such as trigger processors, could be seriously envisaged. These applications are today the most demanding concerning speed of execution. It was recognized quite early (see, e.g., (9)) that pattern recognition tasks, such as event

* Author's e-mail: cmk@mppmu.mpg.de

Table 1. Major developments in neuromorphic hardware. The performance of these multiprocessor chips is indicated by the execution time of a specific multi-layer perceptron. The corresponding network topology is indicated by $i \times h \times o$, where i is the number of inputs, h the number of hidden nodes, and o the number of output nodes. For the digital devices typical execution times are about 5 to 10 μs. The number of bits available for the input data and the weights is indicated. For the accumulation of the neural activity the precision is usually augmented by at least 8 bits.

Analog Devices					
Chip name	manufacturer	year	# of proc.	net topology	exec time
ETANN (3)	Intel Corporation (USA)	1991	64	$64 \times 64 \times 4$	8 μs
NeuroClassifier (4)	Univ. Twente (NL)	1994	6	$70 \times 6 \times 1$	50 ns
Digital Devices					
Chip name	manufacturer	year	# of proc.	clock	precision
CNAPS (5)	Adaptive Solutions (USA)	1993	64	20 MHz	8/16
MA16 (6)	Siemens (D)	1994	16	50 MHz	16/16
Totem (7)	Uni. of Pisa (I)	1994	32	30 MHz	16/8
SAND1 (8)	KfK Karlsruhe (D)	1995	4	50 MHz	16/16

selection at the trigger level, could develop into very fruitful applications of neural networks.

Table 1 gives an overview of the major developments in the field of dedicated neural processor systems. The first large-scale industrial production came from Intel Corporation (3), the famous *E*lectrically *T*rainable *A*rtificial *N*eural *N*etwork ("ETANN") chip. This chip was of the analog variety and was able to carry out in parallel 64 multiply/add operations in less than 5 μs. The weights were stored on chip in an analog way with a precision of roughly 4 bits. The ETANN was used first in a test experiment (10) at Fermilab, where a rather involved neural approach was realized: From the space coordinates in a three-layer driftchamber, originating from muons leaving a beam dump, two parameters in 16 bit precision were requested from a $12 \times 64 \times 64$ network, namely the origin and the direction of the muon. Since the driftchamber can only provide coordinates in one plane, the two parameters translate into intercept and slope of the particle trajectory projected onto the plane transversely to the incident beam. The drift time was converted to a voltage as analog input to the ETANN. Each output node of the network was trained offline to represent a bit of each of the two 32-bit quantities. Compared to standard track reconstruction the network behaved extremely well (10), reaching the offline results within a factor of 2 in precision. The main strength of this test experiment, however, was to show that a particle trajectory could be determined with good accuracy at the few microsecond scale, so that serious trigger applications could be imagined.

Another approach, also tried out at Fermilab, was a more classical pattern recognition task, namely shower shape evaluation, discriminating electrons from hadrons in the electromagnetic calorimeter of the CDF detector (11). Here, the network was of a constructed type, rather than trained, using the 5 by 5 array of calorimeter cells around the shower maximum as inputs. The weights were calculated based on isolation and energy sharing criteria in the array. The intention of the neural approach was to trigger on electrons coming from semileptonic decays of B-mesons.

Although demonstrating impressively the validity of the neural ideas for triggering, the main drawback of these approaches, finally preventing any further trigger development, was the delicate analog nature of the ETANN chip: It had to be trained in a cumbersome way to maintain its properties (the preloaded weights, thresholds) over time and it furthermore was suffering from the temperature instabilities inherent to most analog devices.

With the availability of digital neuromorphic chips (see Table 1) a number of real trigger applications emerged: At CERN the WA92 experiment (12) realized an electron trigger for charm events based on the Siemens MA16 chip, replacing the ETANN in their first prototype. In nuclear physics a neutrino oscillation experiment (13, 14) at the Chooz reactor utilized the CNAPS chip as a fast trigger to discriminate neutrino interactions from background cosmic events: The signature of an inverse β-decay $\bar{\nu}_e p \rightarrow e^+ n$ is a prompt signal from the annihilation photons, followed by a delayed large signal from neutron capture with a typical time difference of about 2 μs. The neural network was trained to reconstruct, from a set of digitized PMT time and energy signals, the position and energy of the event within the fiducial volume of the target. The neural trigger was running at the second level, following a first level trigger based on PMT pulse height discrimination. Using the CNAPS, the task is performed in about 150 μs, allowing for much higher rates at the first level and consequently much lower thresholds than the standard Chooz trigger. Several other applications of the neuromorphic processors presented in Table 1 have been reported (see, e.g., (15) and the proceedings of previous conferences of this series (16)).

One should realize, however, that the number of concrete running applications using neuromorphic hardware are still scarce. This is mainly due to the fact that most of the classical large-scale industrial applications, such as online character recognition (OCR) do not require execution times at the microsecond scale. In these fields modern Pentium-type serial computers are fast enough. So it is not too surprising that most neuromorphic developments by industry have died out these days (ETANN, MA16, CNAPS, just to name the most prominent of them). As time progresses, however, and the problems to tackle will become more demanding, either concerning execution time or complexity of the problem, the parallel neuromorphic way seems the only viable solution: In the field of realtime triggering or pattern recognition, neural networks play, due to their inherent parallelism of computation, a unique role concerning speed, discrimination power, and adaptability to highly complex tasks. In this sense we foresee a renaissance of neuromorphic hardware in the future (see, e.g., the Silicon Brain Project (17)).

THE H1 NEURAL NETWORK TRIGGER

In the following we will describe in more detail a very ambitious neural trigger project, operational in the H1 experiment at the electron-proton collider HERA at DESY. The H1 neural network trigger project presently is the largest such application in high energy physics and is running successfully now since a number of years. For the future high luminosity running at HERA an upgrade of the neural preprocessing is underway, which aims at constructing, in hardware, physically motivated "intelligent" variables.

Detector and Trigger Scheme

A modern, large particle detector system such as H1 at HERA is built with the intention to serve as a general purpose facility and be prepared to be sensitive for the expected physics as well as potentially new phenomena. To this end a large variety of detection principles are employed, covering efficient detection and measurement of hadronic particles, photons and leptons (electrons and muons). For charged particle detection H1 uses an arrangement of drift and proportional chambers surrounding the interaction point (central tracking detector) and the forward (proton direction) region (forward tracker). For particle energy measurement the tracker system is surrounded by a central liquid argon calorimeter and two warm calorimeters to cover the extreme forward ("plug") and backward ("SpaCal") parts of the solid angle. The calorimeter is enclosed by a superconducting solenoid with diameter of about 6 m, followed by an instrumented magnetic flux return yoke for muon measurement. In the forward direction, a warm toroidal magnet, sandwiched between sets of drift chambers, allows additional measurement of high energy muons. For photoproduction reactions and for luminosity measurements several electron and photon taggers are installed upstream of the proton beam. In the downstream part Roman pots and a neutron calorimeter are installed. Further details on the H1 detector are given elsewhere (18).

For triggering the apparatus, H1 has installed a scheme of three levels, two hardware levels and one software level ("level 4"). An intermediate software level ("level 3") is provided, but not used at present. At level 1, each of the detector components (subdetectors) provides a set of triggers to a central trigger box, where they can be subjected to simple coincidence logic. Since the HERA bunch cross frequency is 10.4 MHz (corresponding to a time interval of 96 ns between possible interactions), the level 1 triggers must be able to derive a trigger decision each 96 ns. This is achieved by storing the information for each bunch cross (BC) in a digital pipeline for a maximum of 30 BC's. After about 2.3 μs (24 BC's) a level 1 trigger decision is formed and the information is transferred to the level 2 systems. At this point the primary deadtime starts, no further triggers can be accepted until a fast clear signal from the level 2 trigger system has been issued rejecting the event. When the event is accepted by the level 2, the detector readout is initiated and the full event information is sent to the level 4 processor farm, where a full event reconstruction is performed and the final event decision is taken.

For the level 2 hardware trigger the decision time is limited to 20 μs in order to digest a maximum of 1-2 kHz from level 1 while keeping the deadtime below 2 %. At level 2, the information from all level 1 processors is available, so that "intelligent" use of this information is possible, exploiting the correlations among the various trigger quantities. The output of the level 2 trigger must not exceed 100 Hz which is the maximum rate for the level 4 RISC processor farm. The output rate of level 4 is dumped to permanent tape and is limited to about 10 Hz.

The Need for Complex Triggers

Since its start in 1992 the HERA *ep* collider has constantly improved its performance. Already during the first years of operation it became clear that many high cross section physics reactions, most importantly low Q^2 processes and photoproduction could not be triggered with their full rate due to the competing background (dominated by beam-gas reactions). More severe level 1 conditions could not help out due to substantial loss in physics efficiency. The "natural" way out was to scale down ("prescale"), sometimes quite heavily, the respective level 1 triggers designed for these physics reactions.

In 1995 H1 has tried to improve on this unsatisfactory situation by commissioning two hardware systems at the second trigger level, one of them a novel digital neural network trigger based on the CNAPS chip. The level 2 triggers have access to the information provided by the various subdetectors at level 1 and can use this information in an "intelligent" way.

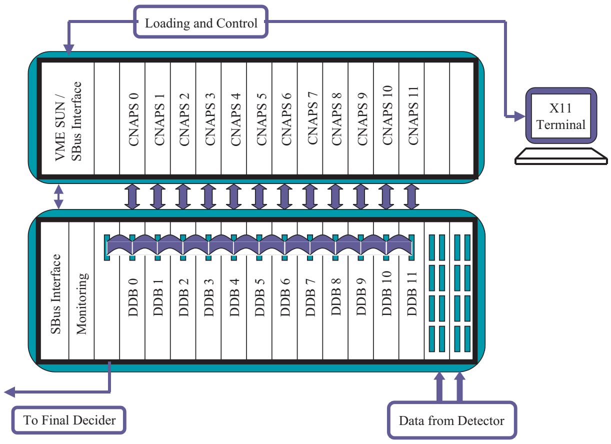

FIGURE 1. *Layout of the H1 Neural Network Trigger system ("L2NN"). Each network processor (CNAPS) is associated with a preprocessing module (DDB) preparing the net inputs individually for its companion CNAPS board. The system is steered by a VME SPARCstation with remote accessibility.*

Architecture of the Neural Network Trigger

The level 2 trigger system has the task to be able to discriminate all possible physics reactions. This request has a very important implication for the neural architecture in the H1 level 2 neural trigger system: Our investigations have shown that *small* nets trained for *specific* physics reactions, working all in parallel, are more efficient compared to a single larger net trained on all possible physics reactions simultaneously. Most importantly, putting these nets to a real trigger application, the degree of modularity is extremely helpful when a new trigger for a new kind of physics reaction is to be implemented: there is no need to retrain the other nets, the new physics net is simply added to the group of the others.

For the network computations (matrix-vector multiplication and accumulation) the CNAPS chip by Adaptive Solutions (5) was used. For the preparation of the input quantities and their interfacing to the CNAPS chip dedicated hardware has been built at the Max-Planck-Institute (preprocessing hardware, see below). Due to the high flexibility of programming the CNAPS chip, arbitrary algorithms can be realized, provided they fit in the latency of level 2 (L2). In the L2 Neural Network Trigger of H1 three different algorithms are used at present:

- Feed Forward Networks: these networks have a fully connected three-layer structure with one input and one hidden layer (with a maximum of 64 nodes each), and one output node. The input layer is fed with the components of a vector $\vec{x}$ spanning the "trigger space", prepared by our preprocessing hardware. The output y is used as a discriminator to make the trigger decision. Details on the network architectures chosen and on the methods to determine the weights $\vec{w}$ are found elsewhere (19).

- Background encapsulators: To evade possible bias in selecting a specific physics class for training against the background, various approaches such as self-organizing networks or radial basis function networks for encapsulating the background are under study.

- Constructed Nets: For some simple, low-complexity applications a topological correlator, exploiting the fast matrix-vector multiplication hardware of the CNAPS chip, is used (20).

The strategy of using the networks is the following: Each of the networks is trained for a specific physics channel and is coupled to a set of level 1 subtriggers, particularly efficient for that channel. Because the level 1 subtriggers are sufficiently relaxed to be efficient, their rate is usually unacceptably high. The level 2 trigger therefore has the task to reduce the excess background rate in the subtrigger set while keeping the efficiency for the chosen physics channel high. At present, 12 networks are running in parallel, mostly optimized for production of vector mesons, which are difficult to separate from the background at level 1. Typical rate reductions are between a factor of 5 to over a few hundred.

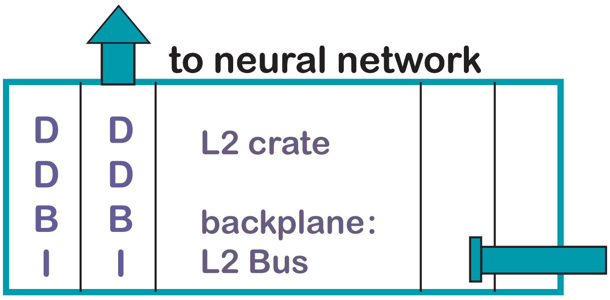

FIGURE 2. *Structure of the DDB-crate which manages the data stream from the level 1 trigger processors to the neural networks: The cable at the right signifies the data from the various subdetectors, which are distributed over a 128bit wide backplane ("L2 bus"). The DDBs pick up the information suited for their respective companion net, prepare the preprocessed net input quantities, and send them to the CNAPS boards.*

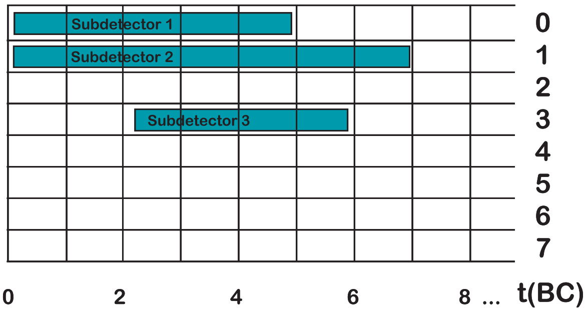

FIGURE 3. *Details of the data flow on the backplane of the DDB-crate: The data from each subdetector are transmitted as a sequence of 16 bit words on one of the 8 subbuses of the L2 bus (gray areas). In the figure three subdetector data streams are shown. The total amount of subdetectors supplying data for L2 is eight at present. The DDBs are programmed to pick up the respective set of items from the subbuses foreseen for preprocessing for the companion neural network.*

The Trigger Hardware

According to the principles described above, the hardware realization for the network trigger is chosen (see Fig. 1): Receiver cards collect the incoming trigger information of the various subdetectors and distribute them via a 128 *bit* wide L2 bus to the preprocessing units, called *Data Distribution Boards* (DDB). Each DDB is able to pick up a freely choosable set of items from the L2 data stream. It performs some basic operations on the items (e.g bit masking, summing) and provides an input vector of maximally 64 8*bit* words for one CNAPS/VME board. Controlling and configuring of the complete system is done by a THEMIS VME SPARCstation, which is located in the crate housing the CNAPS/VME boards.

The CNAPS board

The algorithms calculating the trigger decision are implemented on a VME board housing the CNAPS chip, which is a parallel fixed-point computer in SIMD architecture. The CNAPS-1064 chip (also called *array*) houses 64 processor nodes (PN). Up to eight chips (512 PN's in total) can be combined on one board. A PN is a processor for itself except that it shares the instruction unit and I/O buses with all other PNs. The instruction unit, the sequencer chip CSC-2, is responsible for the command and data flow. The commands are distributed via a 32 *bit* PN command bus. The 8 *bit* wide input and output buses are used for the data transfer to and from the CNAPS arrays. A direct access to these I/O buses is realized with a mezzanine board developed by us. Through this board the input vector is loaded into the CNAPS chip and the trigger result is sent back to the DDB without significant time delay. For synchronization reasons the CNAPS boards are driven with an external clock at 20.8 MHz (2 times the HERA clock frequency of 10.4 MHz).

The main internal parts of the PNs are arithmetic units like adder(32 *bit*) and multiplier(24 *bit*), logic unit, register unit, 4K memory and a buffer unit. All units are connected via two 16 *bit* buses. Calculations are done in fixed-point arithmetic. The sigmoidal transfer function is implemented via a 10 *bit* look-up table. A feed forward step with 64 inputs, 64 hidden nodes and 1 output node can be computed in 8 μs at 20.8 MHz, or in 166 clock cycles. To get the same result with a single conventional CPU one would have to clock it at several GHz.

The Data Distribution Board(DDB)

The Data Distribution Board resides in a special *"L2 VME crate"* equipped with the L2 Bus, an 8 times 16 bit parallel data bus (see figs. 2, 3) running with the HERA clock in an interleaved mode, yielding an effective 20 MHz transfer rate. For each subdetector the level 1 data are sent serially onto one of the eight subbuses of the L2 backplane. For system control purposes, a special monitor board ("spy") with an independent readout of the data transmitted over the L2 bus is residing in the same crate. The data is a heterogeneous mixture from different subdetectors, e.g., calorimetric energy sums, tracker vertex histograms, tracker rays (bits in the $\theta - \phi$ plane), bit-coded muon hit maps, etc.

On the DDB, the L2 data received are passed through a data type selection where they can be transformed (e. g. split into bytes or single bits) using look-up tables (*LUT*).

After bit splitting, several preprocessing algorithms like summing of bits and bytes, bit selections or functions (again realized via LUT's) can be applied. The data may also be sent unchanged to a selection RAM, where the input vector for the neural network computer is stored. Through the use of XILINX 40XX chips the hardware can be flexibly adapted to changes, e.g., for new data formats in the received input. Using selection masks the data are transmitted via a parallel data bus to a mezzanine receiver card directly connected to the local data bus on the CNAPS board.

The complete system with the 2 VME crate is shown in Fig. 4, as it is installed in the H1 experiment (on top the CNAPS crate with the VME control computer, below the DDB crate housing the L2 data bus). Further details on the neural trigger system can be found elsewhere (21).

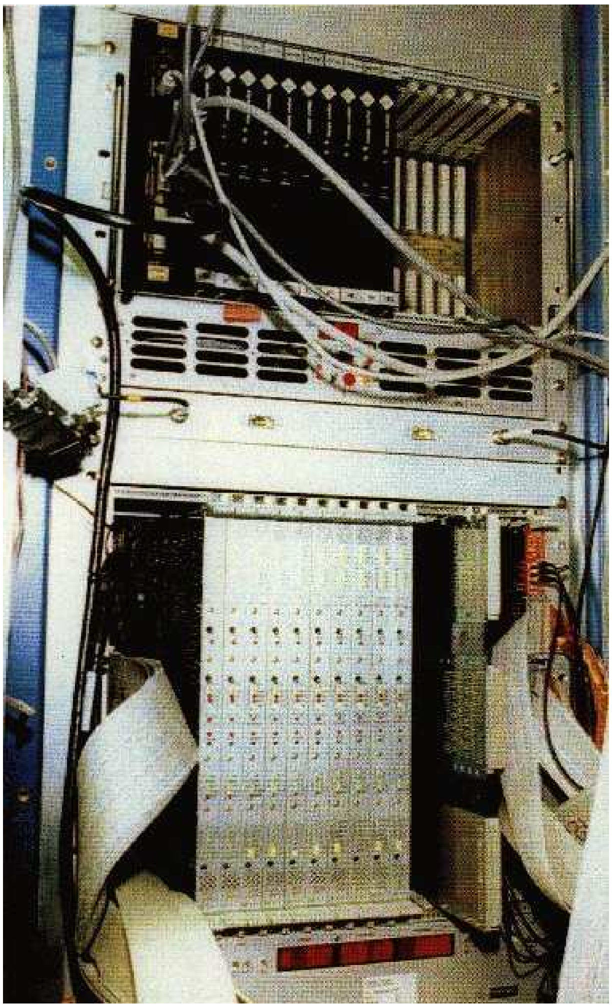

FIGURE 4. *View of the H1 Neural Network Trigger system (L2NN). Each network processor (CNAPS, upper crate) is connected with a preprocessing module (DDB, lower crate). The cables at the lower right come from the subdetectors, the ones at the lower left carry the decision bits from the neural networks.*

SOME PHYSICS RESULTS WITH THE NEURAL TRIGGER

With the setup described above various physics channels have been investigated, most importantly elastic and inelastic vector meson photoproduction (ρ, ϕ and J/ψ and Υ). These reactions are difficult to trigger on efficiently since their event signatures are very close to the high rate background (few low momentum tracks in the main detector). Figure 5 shows the trigger rates into 10 different neural networks (red lines), the output rates (blue lines and the rate reduction factor (green lines) for a typical HERA day. As example, for J/ψ elastic photoproduction three different nets are installed (box 2, 4 and 6), depending on the photon-proton center of mass energy with correspondingly quite different detector signatures. Typical rate reductions are between a factor of 5 (network 2) and over 200 (network 6). With these networks a sizeable sample of elastic J/ψ photoproduction events was triggered and led to published cross section measurements shown in Fig. 6. Most remarkable is the fact that the cross section measurements, with the help of the neural trigger technique, could be extended to the highest HERA center of mass energies, where so far no data have existed. An important physics message is extracted from these data: The standard Regge expectation is not in agreement with the data, in contrast to the QCD prediction. The validity of the QCD calculations rely on the fact that the J/ψ meson is heavy enough to justify the perturbative approach.

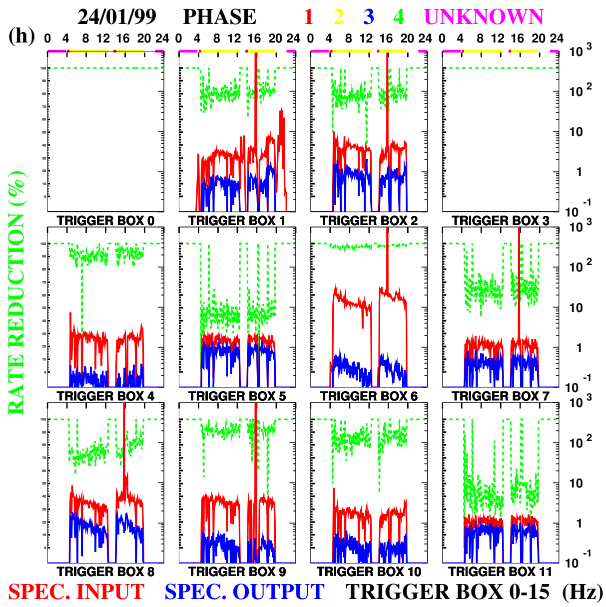

FIGURE 5. *Online rate measurements for the twelve neural trigger boxes operational during the run period of early 1999 (see text).*

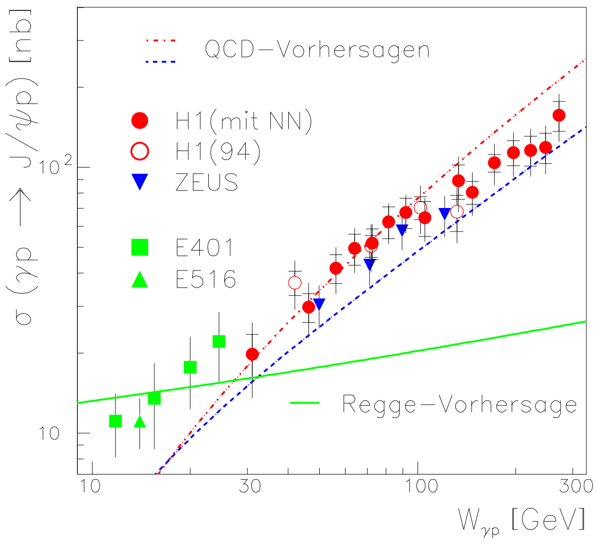

FIGURE 6. *Measurements of the total cross section for exclusive photoproduction of J/ψ mesons as function of the photon-proton center of mass energy W. The H1 points are obtained with triggers from three different neural networks, depending on the first level triggers and the observed final state topology. The measurements are compared to those from ZEUS and to some fixed-target experiments at lower energy from Fermilab.*

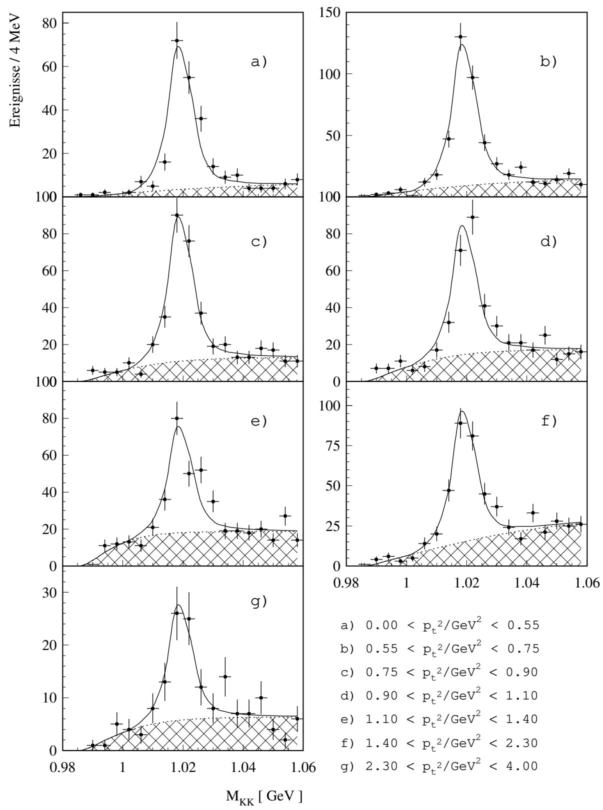

FIGURE 7. *Invariant mass distributions from the two final state particles in events from elastic photoproduction of ϕ-mesons, triggered by a neural network trigger. The various distributions correspond to different bins in transverse momentum of the kaon pair. Clear peaks at the ϕ mass are observed.*

Another example, close to the final analysis, is the photoproduction of ϕ mesons (see the spectra from the K^+K^- invariant mass spectrum in Fig. 7). This channel could only be triggered with the help of the neural technique.

NEW PREPROCESSING ALGORITHMS: THE DDB2

HERA will move on to improve its luminosity by a factor of 5 in the year 2001. This increase will be achieved mainly through focusing magnets inserted close to the interaction regions of the collider experiments. Since the output rate to tape is to be limited to less than 10 Hz, the H1 trigger system will face the challenge of increased rejection power compared to the present situation. As a consequence of this a new preprocessing hardware is being built for the neural network trigger with the aim to provide the network with more physics-oriented quantities.

The idea behind the improved preprocessing is that the "trivial" part of the correlations in the trigger data are determined, namely the association of information from the various subdetectors linked by the passage of particles or jets, as sketched in Fig. 8. In this way physical objects are formed, defined by their topological vicinity in the various parts of the detector. With the new preprocessing the full granularity of the level 1 trigger information is exploited, but the input data volume to the networks is limited to the physically relevant information.

The central algorithm of the preprocessing is clustering in the various subdetectors, with subsequent matching in the angular coordinates θ and ϕ. To be specific, the following steps will be performed in dedicated hardware, making extensive use of fast modern FPGA technology (for a more detailed discussion see (22)):

Cluster Algorithms: The energy depositions in both the electromagnetic and the hadronic parts of the LAr calorimeter will be clustered, using the highest available granularity (656 "trigger towers" or TTt's), summing nearest neighbors around a local maximum. Before the clustering, look-up tables will perform arbitrary transformations on the TT energies (e.g., transverse energy, taking into account the polar angle θ). Double counting of touching towers is explicitely avoided by the algorithm.

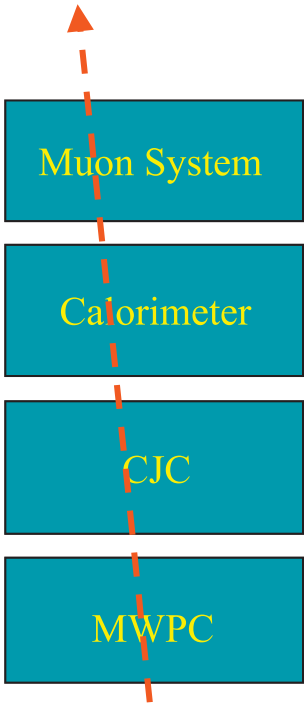

FIGURE 8. *Principle of the preprocessing algorithm for the DDB 2: Particles or jets traverse the various subdetectors of the H1 experiment. The clusters found in each subdetector are correlated in the angular variables* θ *and* ϕ.

For each cluster found, the total energy $Etot$, the center $Ecent$ with its $\theta-\phi$ values, the ring $Ering$ and the number *nhit* of towers containing energy above a suitable threshold are stored.

Bit fields (such as hit maps from the MWPCs, the drift chambers or the trigger cells of the SpaCal calorimeter) will also be subjected to a cluster algorithm: In this case a preclustering will be performed, summing all immediate neighbors to a given "seed" bit. After this preclustering, which results in a "hilly" $\theta-\phi$ plane, the same algorithm as for the calorimeter is executed.

Due to their coarse granularity, the trigger information from the muon chambers will not be clustered, but the $\theta-\phi$ information of the muon hits is stored.

Matching: Based on common angular coordinates (θ and ϕ) the clusters from the various layers of the detector will be gathered in physical objects, forming a vector the components of which represent the list of cluster quantities determined in the previous clustering step. In addition, the objects will be ordered in magnitude according to chosen components. Three parallel sorting machines are foreseen, delivering arrays of vectors sorted according to three pre-determined vector components. Ordered input was found very helpful in speeding up the learning and in improving the discrimination power.

Net Inputs: The input to the networks are the sorted vector components (or a freely choosable subset, depending on physics reaction considered). Due to the serial clocking-in of data into the CNAPS chip, a limit of about 8 to 10 objects (with 8 components each) as net input is imposed, which should be quite sufficient for the physics applications considered at present, such as heavy flavor production, both open and hidden, jets in photoproduction, low Q^2 reactions and other physics channels depositing mostly low energy in the calorimeter.

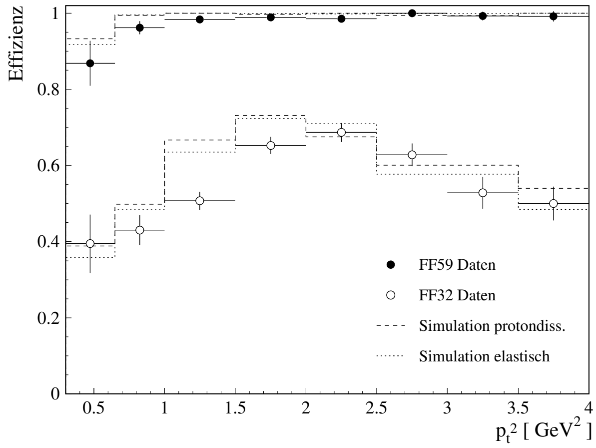

FIGURE 9. *Efficiency for a network trained to select elastic photoproduction of* ϕ*-mesons, using inputs obtained from the present DDB compared to the efficiency with quantities supplied by the DDB 2 (upgraded preprocessing). The improved efficiency by the DDB 2 is obvious (24).*

Post processing: A final step is considered which determines some physical quantities from the vector components, such as cluster counting, angular differences etc. The exact specifications for the post-processing is still under investigation, studying specific physics reactions.

As an example for the increased selection power of the DDB 2 with respect to the less sophisticated DDB ("DDB 1") presently operating, the elastic photoproduction of ϕ-mesons has been studied using the neural network trigger. This reaction could be observed in H1 for the first time (23). The performance of the network based on DDB 1 quantities is shown in Fig. 9. Using in a simulation the new physics objects (24) provided by the DDB 2, the selection efficiency - at constant background rejection - could be increased by more than 40 percent (see Fig. 9).

CONCLUSIONS

We have presented the principles and the hardware realization for a second level neural network trigger, operational in the H1 experiment at HERA as a global event decision machine since summer of 1996. Based on commercially available, massively parallel digital ULSI neural network chips, a 20 μs decision time is achieved for the network trigger. The network inputs are derived from the trigger information provided by the various level 1 trigger systems and are preprocessed by custom-designed hardware (DDB 1). The system has demonstrated its physics potential by largely increasing the statistics for vector meson production, producing the first physics result from a neural network trigger. The neural trigger also allowed observing elastic ϕ photoproduction for the first time in H1.

For the data taking with the upgraded HERA machine higher selectivity of the neural trigger is needed, and an improved preprocessing unit is under development (DDB 2), performing a quasi physics analysis at the trigger level: Physical objects ("particles" , "jets") are created from the full granularity of the level 1 trigger processors, extracting the physically relevant information while keeping the data volume into the networks manageable. The proposed unit was subjected to test reactions, demonstrating its superior selection power compared to the present DDB 1.

The fact that industry has discontinued the production of neuromorphic hardware is not a sign of lost interest in neural network realtime applications. It is just a reaction to the market which in most of its applications is satisfied with a time response within fractions of seconds, easily obtainable with present-day Pentium type computers. For the visible future realtime applications in physics experiments (high energy physics, astrophysics etc.) trigger processors will have to provide response times in the microsecond range, only obtainable with parallel computing architectures. Developments in this area are underway in the academic world (see, e.g., (25)). We are convinced that these efforts will not only be beneficial for science but also leave their impact for future commercial applications.

REFERENCES

1. P. Bhat, "Advanced Analysis Techniques", Invited Talk, these proceedings.
2. H. Prosper, "Artificial Intelligence - Offline Applications", rapporteur talk, these proceedings.
3. ETANN 80170NX Chip, Specification Booklet, Intel Corporation, 1991.
4. Neuro-Classifier Chip: P. Masa and H. Wallinga, University of Twente, Internet: www.ice.el.utwente.nl/Finished/Neuro/NeuroClassifier.htm
5. CNAPS Release Notes 2.0, Adaptive Solutions, Inc. , Beaverton Or. (1993).
6. U. Ramacher et al., "Design of a First Generation Neurocomputer", in: VLSI Design of Neural Networks, eds. U. Ramacher and U. Rückert, Kluwer Academic Publishers, 1991.
7. G. Anzellotti et al., in "New Computing Techniques in Physics Research", eds. B. Denby & D. Perret-Gallix, World Scientific, Singapore 1995.
8. SAND1 chip, Forschungszentrum Karlsruhe and Institut für Mikroelektronik, Stuttgart. Internet: fuzzy.fzk.de/ fischer/sand/index.html
9. B. Denby, Comp. Phys. Comm. 49 (1988) 429.
10. C. Lindsay et al., Nucl. Instr. Meth. A317 (1992) 346.
11. M. Badgett et al., "A Neural Network Calorimeter Trigger Used in CDF", Proceedings of the 1992 IEEE Nuclear Science Symposium, Orlando, Florida.
12. A. Baldanza et al., in "New Computing Techniques in Physics Research III", eds. K.-H. Becks, D. Perret-Gallix, World Scientific, Singapore, 1994.
13. M. Apollonio et al., Phys. Lett. B420 (1998) 397
14. A. Baldini et al., Nucl. Inst. Meth. A389 (1997) 144
15. B. Denby, Comp. Phys. Comm. 119 (1999) 219.
16. Proceedings of the Workshop on New Computing Techniques in Physics Research IV, eds. B. Denby, D. Perret-Gallix, World Scientific, Singapore, 1995.
17. The Silicon Brain Initiative, Irvine Sensors Inc., Internet: www.irvine-sensors.com
18. I. Abt et al. "The H1 detector at HERA", Nuclear Instruments and Methods in Physics Research A 386 (1997) 310.
19. S. Udluft et al., in "New Computing Techniques in Physics Research VI", eds. G. Athanasiu & D. Perret-Gallix, Parisianou, Athens, 2000, p. 166.
20. J. Möck, "Untersuchung diffraktiver J/Ψ-Ereignisse im H1-Experiment bei HERA und Entwicklung neuronaler Triggeralgorithmen", PhD thesis, TU München (1997).
21. J.H. Köhne et al., Nuclear Instruments and Methods in Physics Research A 389 (1997) 128.
22. J.-C. Prevotet et al., "Intelligent preprocessing for Neural Networks in the H1 experiment", these proceedings.
23. F. Gaede, *Exklusive Produktion von ϕ-Mesonen in ep-Streuung am H1-Experiment bei HERA*, PhD thesis, Christian-Albrechts-Universität zu Kiel (1997)
24. S. Udluft, *Protondissoziative Photoproduktion von ϕ-Mesonen am H1-Experiment bei HERA*, PhD thesis, Ludwig-Maximilians-Universität, München (2000)
25. B. Granado, L. Lacassagne and P. Garda, "Maharadja: A candidate system for RBF networks with the Mahalanobis distance in HEP", in "New Computing Techniques in Physics Research VI", eds. G. Athanasiu & D. Perret-Gallix, Parisianou, Athens, 2000, p. 80.