Parallel Processing in Neural Systems and Computers
R. Eckmiller, G. Hartmann and G. Hauske (Editors)

Parallel Process Interfaces to Knowledge Systems

K.H. Becks†, W. Burgard*, A.B. Cremers*, A. Hemker† and A. Ultsch*

*Department of Computer Science
University of Dortmund
D-4600 Dortmund 50

†Department of Physics
University of Wuppertal
D-5600 Wuppertal 1

This paper deals with the integration and combination of connectionist approaches and symbolic knowledge based systems. Our aim is the development of expert systems which are tightly coupled to real world processes. We discuss architectures for such systems and present two possible outlines of an integrated system for the analysis of data produced by a high energy physics experiment.

1. Introduction

The wide field of domains in which expert systems solve complex problems such as diagnosis, construction, planning, scheduling etc. shows the success of the symbolic knowledge representation paradigm. One expert system which performs fault diagnosis in a high energy physics experiment is presented in [Beck89]. Symbolic knowledge representation has a clear advantage: the knowledge chunks are communicable i.e. they can be explained, taught, learned and easily modified. On the other hand, the symbolic approach lacks of inherent methods to refine raw incoming data into structured symbolic information and to tackle with inconsistencies, such as erroneous data, missing values, and unexpected events. Both properties, for example, are required if expert systems have to be connected to complex external processes in order to solve tasks such as data analysis or control. In this case a further important task which has to be carried out is the gathering of sensory information, the influence of the sensors, the processing and filtering of the data in an appropriate and possibly adaptive manner.

Connectionist models, however, claim to capture more precisely more complex and imprecise properties of the real world. They are inherently parallel and furthermore adaptive in the sense that knowledge can be learned directly from experience.

Hence, the combination of symbolic and connectionist approaches or the use of integrated systems seems to be sensible if expert systems have to solve tasks which require a tight coupling to the environment they are embedded in. Different approaches to achieve an integration of expert systems into given environments will be discussed in the following section. For illustration we concentrate on the requirements for such systems in the framework of a high energy physics experiment.

2. Combination of Connectionist and Symbolic Systems

First, connectionist models can be an integral part of an expert system. They can be used to represent the rules of a system (see [Gall88]) or to implement a fuzzy reasoning calculus (see [Beck87]). Furthermore they can be integrated into the reasoning process itself. In the PANDA system (see [Ults89a, Ults89b]) connectionist models are used to learn symbolic rules from given data and to observe the reasoning process. In this system a neural network "learns" structural properties of proofs to the effect that the same proof or similar proofs can be performed more efficiently the next time.

Second, a connectionist system can operate as a sub-system which condensates a high volume of data describing complex states to information which is needed as input for an expert system.

Third, both paradigms can be integrated into one. This was carried out with the development of classifier systems which originally were suggested by John Holland (see [Holl76]) as an outgrowth of his work on genetic algorithms. So far, mainly the machine learning society was interested in classifier systems because they perform a kind of data intensive learning by example, with sparse reinforcement. Although classifier systems are rule based, they work on a subsymbolic level (see [Gold89, Davi87]). The basic idea of classifier systems should enable them to unify the distinct cognitive abilities, which are required to solve a complex problem. On the one hand classifier systems are capable of communicating directly with the environment which enables them to work on low-level tasks, on the other hand symbol-level activities can also be implemented (see Forr85]). Rule-learning also covers the whole range from signal processing rules to problem solving rules, although the internal representation scheme is always subsymbolic.

Because of their ability to interact directly with the environment and their adaptive behaviour, classifier systems are capable of dealing with dynamic environments. All parts of a classifier system are inherently parallel. There are applications of classifier systems in quite different domains ranging from poker playing, gas pipeline control to simulation of an "animal-like automaton". So far classifier systems did not succeed in larger applications, although no principal limitations have yet been encountered.

3. The DELPHI Experiment

The new electron positron collider LEP at the European High Energy Physics Research Center CERN near Geneva has started operating in July '89. In the collider ring electrons and their anti-particles, the positrons, are accelerated in opposite directions. Each time an electron collides with a positron, in a so called event, many new and partially short-living particles arise. Some particles can be identified with the help of large detector systems, one of which is the DELPHI detector. In order to show the difficulties in analyzing such an event we subsequently will give a short and simplified description of the involved physical processes (for more details we refer to [Davi89]).

The electron-positron annihilation process produces an initial quark-antiquark pair and possibly several gluons. Quarks are elementary particles and are the fundamental constituents of matter. There are six known varieties of quarks distinguished by their flavours: up, down, charm, strange, bottom, top. The top quark is of major interest in the DELPHI experiment.

After the creation of the initial quark-antiquark pair new quark-antiquark pairs with less energy are generated in a fragmentation phase. Fragmentation is a cascading process in which each quark-antiquark pair can produce new pairs with less energy. Quark and anti-quark always have the same flavour e.g. up and anti-up. Quarks of different flavours cluster and form new particles (hadrons). Some produced short-living hadrons decay into long-living particles according to known rules. Most of these stable particles are observed with the help of particle detector systems (see figure 1).

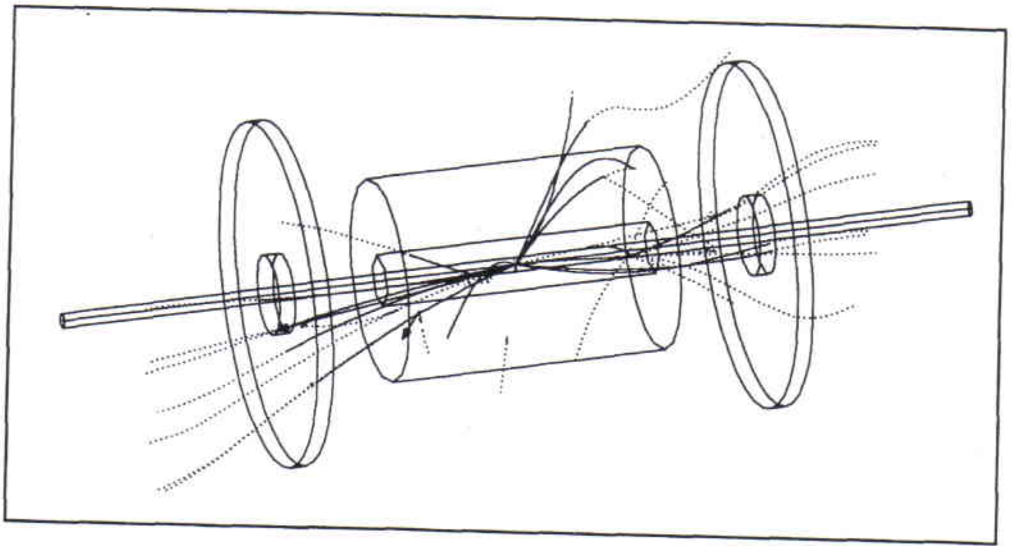

Figure 1: Typical event observed by parts of the DELPHI detector

Typical questions for an event concerning the intermediate states of the physical process are:

- What was the flavour of the initial quarks?
- Were there any exotic particles e.g. Higgs or rare decays?
- How many gluons were emitted resp. how many jets occurred?

It is very difficult to answer these questions because the whole physical process is extremely non-deterministic and the detectors can observe the physical processes only with a limited efficiency.

The fragmentation process itself is not directly observable as well as many decay processes of hadrons. It is very difficult to detect some of the neutral particles and furthermore noise and blind sections in the detector can inhibit the observation and identification of stable particles.

The non-determinism of the physical process has several reasons. The flavours of the produced quark-antiquark pairs are randomly distributed and the depth of the fragmentation cascade is unknown. Although the quark constituents of a hadron are known, it is impossible to observe the quarks which cluster to a given hadron. A further source of uncertainty is that different decay rules are possible for an unstable hadron.

4. Analysis Strategies

The average number of observed charged particles per event is about forty. Inferring back from these observed particles to the preceding physical process expands an extraordinary complex search space. Nowadays most of the physics analysis programs avoid this problem by looking only at the signature of the final annihilation state to answer questions about an event. The huge amount of data from each event (the number of events is 1-2 per second where each event has an average data size of 150 Kbyte) is reduced to only a few physical, statistical and geometrical parameters which are sensitive to a certain question. It is left to the intuition of the physicist to find out the relevant parameters.

Simulation programs can prove whether a value is sensitive to a certain question. These so called Monte Carlo programs contain the actual knowledge of the elementary physical processes, which could have happened after the collision of electron and positron. With an additional simulation of the detector system it is possible to produce data before starting the real experiment. The simulated data can be used to test and tune the physics analysis programs. Moreover, it is possible to check the underlying physical model by comparing simulated data with real data from the experiment.

Data reduction and flat, statistical based reasoning imply a loss of information. Unusual effects which did not manifest in the selected parameters of the analysis programs cannot be recognized. Differences between simulated and real data cannot be directly interpreted, little deviations are not conspicuous. An ideal system would be able to reconstruct each event using the underlying physical model or would locate the point where an event, that contains new physics, does not match the model. If one attempts to develop an intelligent assistant for the analysis of the event data, one has to build tools which allow a fast condensation of the incoming data. Based on possible architectures for an integration of different paradigms provided in section 2, we now will present an outline of two possible systems which should be able to solve this task. Since this project is still in the starting phase, we only can give proposals for both solutions.

5. Outlines of an Integrated System for Data Analysis

Fragmentation and particle decay suggests the representation by production rules. With these rules contained in the production memory it is possible to reconstruct the event history by abduction. The initial state is equal to the final physical state and the goal state corresponds to the original electron-positron annihilation state. State transitions are defined by applying fragmentation and decay rules backwards. In this case not the goal itself is interesting but the path leading from the initial state to the goal state. An expert system, which is able to reconstruct the physical processes, i.e. to find this solution path in the problem space, could answer all the questions outlined in the previous section.

The non-determinism of the physical processes requires probabilistic reasoning. Since we will have more than one possible solutions in most cases, there has to be a mechanism to judge the quality of solutions. Heuristics to simplify the search are hardly known or not explicitly formulated. Hence the system must be able to learn such heuristics. Monte Carlo simulation can provide learning examples and the protocol of the simulation can be used to control the learning process and to judge solutions.

As pointed out in section 2 a connectionist system could be used to preprocess the incoming data from the detector. This sub-system would benefit from the advantages of the connectionist approach with regard to suppression of detector noise, parallel processing and condensation of quantitative data to qualitative data. The preprocessed data would be the input of a symbolic knowledge based system, whose task is to reconstruct the physical process. An advantage of this possible solution is that the learned heuristics could easily be interpreted by the physicists. Furthermore, the rules to guide the search could be implemented as connectionist systems, and neural networks could be used to improve the search. In this case, however, an interpretation of the learned rules will be very difficult.

As second possible solution of the reconstruction problem would be the application of a classifier system. If this solution is preferred, the question arises how the different types of information can be represented. Messages should be able to encode signals from the particle detector, as well as it should be possible to represent an intermediate states of the physical process by one or more messages. Classifiers could be used for signal processing, but also to simulate a fragmentation step or a particle decay. There should be room left in the representation of a classifier to encode a situation, in which the processing rule should be preferably applied. In the training phase classifiers should learn those situations. Starting with messages encoding signals from the particle detector, the classifier system should explore effectively and in parallel the problem space to find a solution. A backtracking strategy is not necessary because of the parallel rule activation in classifier systems. The environment can check the validity of physical laws after each cycle, i.e. for each reconstruction step. In the training phase the environment can evaluate possible reconstruction steps by comparing them with the Monte Carlo protocol.

The inherent learning capability of a classifier system makes it possible to integrate subsequent changes of the physical model into the system only by changing the environment. It still remains a problem to interpret the heuristics learned by the system. An additional analysis is necessary to investigate the characteristics of those situations, in which one processing rule should be preferably applied.

6. Conclusion

In this paper we have presented different strategies to couple knowledge based systems with connectionist systems. We have pointed out how such combinations can be used to improve the current expert system technology. Connectionist models can be used to couple the expert system to its environment. To achieve a fast refinement from data to information they can build parallel interfaces between real world processes and conventional knowledge based systems. They can also be used as an integral part of an expert system. We discussed classifier systems which can be seen as homogeneous integrated systems. We have proposed two possible outlines of an integrated system to be used for data analysis in a high energy physics experiment. As this example shows the development of tools which allow a tight coupling of expert systems with connectionist systems as parallel process interfaces seems to be a very important task.

References

[Beck87] Becker, L.A. and Peng, J., Using Activation Networks for Analogical Ordering of Consideration, IEEE Conf. on Neural Networks, San Diego, CA, (1987) pp. 367-371.

[Beck89] Becks, K.-H. and Cremers, A.B. and Hemker, A., Design of an Expert-System-Shell for Fault-Diagnosis in the DELPHI Experiments, KI 1/89, (Oldenburg Verlag, 1989), in German.

[Davi87] Davis, L., Genetic Algorithms and Simulated Annealing, (Morgan Kaufman, Los Altos 1987).

[Davi89] Davies, P.C.W., The New Physics, (Cambridge University Press 1989).

[Forr85] Forrest, S., Implementing Semantic Network Structures using Classifier Systems, in: in: Grevenstette (ed.), Proceedings of an International Conference on Genetic Algorithms and their Applications, (Carnegie-Mellon University, Pittsburgh, 1985).

[Gall88] Gallant S.I., Connectionist Expert Systems, Communications of the ACM, Vol. 31, (1987), pp. 52-169.

[Gold89] Goldberg, D.E., Genetic Algorithms in Search, Optimization and Machine Learning, (Addison-Wesley 1989).

[Holl76] Holland, J.H., Adaptation, Progress in Theoretical Biology, Vol. 4, (1976) pp. 263-294.

[Robe87] Robertson, G.G. and Riolo, R.L., A Tale of Two Classifier Systems (1987).

[Ults89a] Ultsch, A., PANDA: Prolog and Neural Distributed Architectures, Technical Report, Department of Computer Science, (University of Dortmund, 1989), in German.

[Ults89b] Ultsch, A. and Hannuschka,R. and Hartmann, U. and Weber, V., Learning of Control Knowledge for Symbolic Proofs with Backpropagation Networks, this volume.