Neural Network Support in a Hybrid Case-Based Forecasting System

B. Lees and J. M. Corchado
Applied Computational Intelligence Research Unit
Department of Computing and Information Systems
University of Paisley, High St., Paisley, PA1 2BE, UK
email: lees-ci0@paisley.ac.uk

Abstract

The progress to date is described in an ongoing project in which the aim is to investigate the combination of case-based reasoning and artificial neural networks as a strategy for cooperative problem solving. The paper describes a successful application in which a Radial Basis Function artificial neural network is used for the adaptation of cases, during the reuse phase of the CBR life cycle. The approach is being applied to the problem of real-time oceanographic forecasting and the results obtained so far are presented.

Introduction

Research into artificial intelligence (AI) has produced various problem solving methods which may be applied to give more powerful computer based problem solving capabilities than may be obtained using purely algorithmic methods. Indeed, the reason for the application of an AI approach is very often precisely because the nature of the problem to be addressed is such that no appropriate algorithm is either known or is applicable. For example, the data and/or knowledge pertaining to the problem at hand may be either incomplete or uncertain. Faced with such a situation the task of the AI professional is to choose, from the various AI problem solving approaches available, a method which appears most appropriate to the problem to be solved.

It is in situations where prior experience of solving similar problems is available that case-based reasoning has proved its value as an AI problem solving strategy. But the nature of a complex problem solving situation may be such that there are different aspects of the problem that may best be addressed through the application of several distinct problem solving methodologies. In such situations the application of a hybrid problem solving approach may be appropriate. In the research described in this paper, the focus is on the combination of case-based reasoning (CBR) and artificial neural networks (ANN) as mutually supportive problem solving methods. The original motivation for this work was twofold: (i) to investigate how neural networks may be employed to assist case-based problem solving, and (ii) to employ such a hybrid approach with the aim of extending earlier work on the application of AI methods to the problem of oceanographic forecasting (Lees *et al.*, 1992). The work forms part of a wider, longer term strategy in which the aim is to investigate the feasibility and applicability of a multi-paradigm approach to artificial intelligence problem solving. In an earlier paper (Corchado *et al.*, 1997) the possibility of combining the problem solving components in the form of a set of intelligent agents was explored. However, in this paper, an approach using a single integrated problem solving mechanism is addressed.

The structure of the paper is as follows. First the integration of CBR and ANN problem solving methods is introduced; a brief outline of work elsewhere on the integration of CBR and neural network methods is given. The application of a hybrid neural network case-based approach, using a radial basis function network for case adaptation as a strategy for real-time oceanographic forecasting, is presented. Finally, a summary of the experimental results obtained to date are presented, which indicate that the approach performs favourably in comparison with the use of statistical and neural network forecasting methods in isolation.

Combining CBR and Neural Networks

Case-based reasoning and artificial neural networks are complementary problem solving methods. Case-based reasoning has the potential to provide, by reference to previous learned experiences, problem solving capabilities in situations which defy attempts to obtain a satisfactory solution through the use of the logical, analytical techniques of knowledge-based systems and standard software technologies: for example, when a clear model of the problem domain is unobtainable. Neural networks are able to analyse large quantities of data to establish general patterns and characteristics in situations where rules are not known and, in many cases, can make sense of incomplete or noisy data. Furthermore, whilst neural networks deal easily, and normally, with numeric (and, to some extent, symbolic) data sets, case-based reasoning can also handle more complex symbolic knowledge structures.

Many complex tasks that a human being can perform with apparent ease, for example distinguishing among visual images, patterns or sounds, are not so easily performed by computers using traditional algorithmic, methods. Neural networks have been found to be a more appropriate means of carrying out such tasks (Rumelhart *et al.,* 1996). The premise underlying the research reported in this paper is that, for certain problems, the integration of case-based and connectionist problem solving paradigms may result in a more effective problem solving facility than would be possible with either paradigm in isolation.

Other researchers elsewhere have investigated the integration of CBR and neural networks as a problem solving strategy. There is current interest in the application of hybrid CBR and ANN systems for the purpose of diagnosis. As an example, the use of a fuzzy logic-based neural network in a case-based system for diagnosing symptoms in electronic systems has been proposed by Liu and Yan (1997), the aim being to overcome the problem that descriptions of symptoms are often uncertain and ambiguous. In the domain of medical diagnosis, Reategui *et al.*, (1997) have used an integrated case-based reasoning and neural network approach. The task of the neural network is to generate hypotheses and to guide the CBR mechanism in the search for a similar previous case that supports one of the hypotheses. The model has been used in developing a system for the diagnosis of congenital heart diseases and has been evaluated using two cardiological databases with a total of over two hundred cases. The hybrid system is able to solve problems that cannot be solved by the neural network alone with a sufficient level of accuracy.

An important task in the design of case-based systems is the determination of the features that make up a case and also of ways to index those cases in the case-base for efficient and correct retrieval. Main *et al.* (1996) consider the use of fuzzy feature vectors and neural networks as a means of improving the indexing and retrieval steps in case-based systems.

A neural network has been employed as a basis for calculating a measure of similarity between a new problem case and each stored candidate case (Garcia Lorenzo and Bello Perez, 1996). It is claimed that the neural network provides a mechanism to retrieve cases using information that in other models would require a parallel architecture.

Further examples of research into the integration of case-based reasoning and neural networks include the work of Agre and Koprinska (1996) and also that of Krovvidy and Wee (1993).

In the area of forecasting, the task of predicting future parameter values from a given sequence of states has been addressed by Goodman (1994); however, his approach was based on simulation, rather than on neural networks.

The Forecasting Problem

A hybrid case-based approach to forecasting is being investigated in collaborative work with Plymouth Marine Laboratory (PML). In this research the aim is to develop a methodology for predicting the values of physical parameters (in particular, sea temperature at a given depth) in three dimensions, around a sea going vessel from data acquired in real time, and also from past records of sea temperature (and possibly other oceanographic parameters) surrounding the vessel at some point ahead of the vessel, which will be reached in the immediate future. This information may also then be used to provide a forewarning of an impending oceanographic *front*, i.e. a boundary between different large water masses. The approach builds on the methods and expertise previously developed at Plymouth Marine Laboratory, and, in particular, in previous collaborative work with the University of Paisley into the application of knowledge based methods for the analysis of oceanographic data (Lees *et al.*, 1992).

The problem of forecasting, which is currently being addressed, may be simply stated as follows:

Given: a sequence of data values (which may be obtained either in real-time, or from stored records) relating to some physical parameter

Predict: the value of that parameter at some future point(s) or time(s).

The raw data (on sea temperature, salinity, density and other physical characteristics of the ocean) which are measured in real time by sensors located on the vessel, consist of a number of discrete sampled values of a parameter in a time series. These data values are supplemented by additional data derived from satellite images, which are received weekly. In the present work the parameter used is the temperature of the water mass at a fixed depth. Values are sampled along a single horizontal dimension, thus forming a set of data points.

This data must be pre-processed in order to eliminate noise, to enhance interesting features, to smooth stable areas and to transform the data set into a form which may be represented on an absolute scale. There are several techniques that can be applied to transform the original data set (Corchado, 1995) to reduce noise, sharpen data and aid in the detection of fronts. The approach adopted employs a *Sobel Filter,* the operation of which is based on the idea that local variations, corresponding to edge transitions, occur at a slower rate than those corresponding to noise.

Hybrid CBR - Neural Network System

In order to produce a forecast in real-time of the temperature a certain distance ahead of the vessel, a *problem case* is generated every 2 km. A problem case consists of a sequence of N sampled data values (after suitable filtering and pre-processing) immediately preceding the data value corresponding to the current position of the vessel. A value of 40 for N has been found empirically to produce satisfactory results. The problem case also includes various other numerical values; these include the current geographical location of the vessel and the time and date when the case was recorded.

The set of N data values forms an *input vector*, which is then used to produce a forecast of the ocean temperature, 5 km ahead. In outline, this process is depicted in Figure 1. (Note that, in practice, it is the set of differences, between the temperature T_i at the current point and the temperature at successive earlier points, which are used as the input vector: i.e. $T_i - T_{i-1}$, $T_i - T_{i-2}$ etc.)

The forecasted values are created using a neural network enhanced case-base reasoning system. The CBR mechanism allows the experience recorded in previous forecasting situations to be reused. The role of the neural network lies in the case adaptation process.

The relationships between the processes and components of the hybrid system are illustrated in Figure 2. The cyclic CBR process shown in the figure has been inspired by work of Aamondt and Plaza (1994). The four basic phases in the CBR cycle are shown as ellipses. Superimposed on the fundamental CBR cycle is a cycle of neural network operations during which the network parameters are retrieved from a neural network knowledge base, employed in case adaptation, and then are revised, with their updated values being stored back in the knowledge base. The full cycle of operations of the hybrid system is explained in the following section.

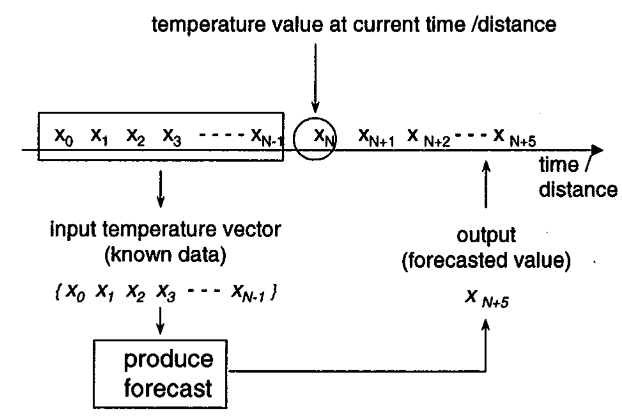

Figure 1 Forecasting operation

The particular type of neural network of interest in the current research is the *Radial Basis Function* (Bishop, 1995), in which the input layer is a receptor for the input data, whilst the hidden layer performs a non-linear transformation from the input space to the hidden layer space. The hidden neurons form a basis for the input vectors; the output neurons merely calculate a linear combination of the hidden neurons' outputs.

Activation is fed forward from the input layer to the hidden layer where a Basis Function is calculated. The weighted sum of the hidden neurons' activations is calculated at the single output neuron. Radial Basis Functions (RBF) are better at interpolating that at extrapolating; where there is a region of the input space with little data, a RBF cannot be expected to approximate well, Furthermore, RBFs are less sensitive to the order in which data is presented to them than is the case with other neural network models, such as Multi-Layer Perceptrons. However, Radial Basis Functions are potentially useful in hybrid systems because of their fast learning capability.

Hybrid System Operation

The forecasting system uses data from two sources: (i) the real-time data are used to create a succession of problem cases, characterising the current forecasting situation; (ii) data derived from satellite images are stored in a database (which, for clarity, is not shown in Figure 2). The satellite image data values are used to generate cases, which are then stored in the case base and subsequently updated during the CBR operation.

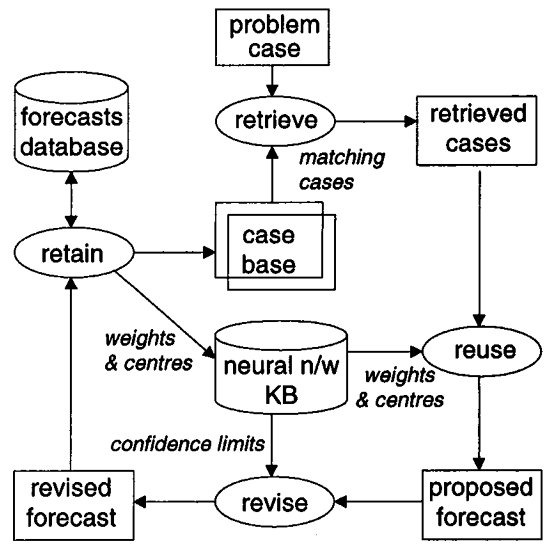

Figure 2 Modified CBR cycle

The cycle of forecasting operations (which is repeated every 2 km) proceeds as follows. (Note that space in this short paper does not permit all the finer details of this process to be included).

First a new problem case is created from the pre-processed real-time data.

A set k cases, which most closely match this current problem case, is then obtained from the case base during the CBR *retrieve* phase, using nearest neighbour matching.

In the *reuse* phase, the values of the weights and centres of the RBF neural network used in the previous forecast are retrieved from the neural network knowledge base. These network parameters together with the *k* closest matching cases are then used to create a forecast of the temperature 5 km ahead. At this point the parameters of the network are modified by taking into account the information contained in the retrieved cases. The effect of this is to allow the system to learn from all these *k* cases (rather than simply using the single adjudged closest matching case) in making a new forecast.

During each forecasting cycle the RBF network is retrained, using the retrieved weights and centres, with the input vectors contained in the *k* matching cases applied as inputs to the network. This process adapts the RBF network, by accommodating the retrieved cases, thus updating the values of the network parameters (empirically, a value for *k* of between 500 to 1000 has been found to be appropriate).

The input vector from the problem case is then fed into the trained network to produce a proposed forecast.

In the *revise* phase, the proposed forecast is modified by taking into account the accuracy of the previous forecasts, which have been reused to obtain the new forecast. Each case has associated with it a cumulative average error which is a measure of the average error in the previous forecasts for which that particular case was used to train the neural network. Confidence limits are calculated by averaging the cumulative average error of the *k* cases used to train the ANN in producing the current forecast. The revised forecast is then expressed, using the confidence limits, as an interval, between upper and lower limits, rather than as a single value.

The revised forecast is then retained in a temporary store – the forecasts database. When the vessel has travelled a further 5 km, the actual value of the water temperature at that point is measured. The forecasted value for the temperature at this point can then be evaluated, by comparison of the actual and forecasted values, and the error obtained. A new case, corresponding to this forecasting operation is then entered in the case base. Knowledge of the forecasting error is also, at this point, used to update the cumulative average error of all the *k* cases that were reused to obtain that forecast.

Radial Basis Function Operation

The RBF network uses 9 input neurons, between 20 and 35 neurons in the hidden layer and 1 neuron in the output layer. Input vectors (explained earlier) form the input to the network; the output of the network is the difference between the temperature at the present point and the temperature a fixed distance ahead. Initially, twenty vectors are randomly chosen from the first training data set and used as centres in the middle layer of the RBF network. All the centres are associated with a Gaussian function, the width of which, for all the functions, is set to the mean value of the Euclidean distance between the two centres that are separated the most from each other.

Training of the network is done by presenting pairs of corresponding input and desired output vectors. After an input vector activates every Gaussian unit the activations are propagated forward through the weighted connections to the output units which sum all incoming signals. The comparison of actual and desired output values enables the mean square error (the quantity to be minimised) to be calculated.

The closest centre to each particular input vector is moved toward the input vector by a percentage α of the present distance between them. By using this technique the centres are positioned close to the highest densities of the input vector data set. The aim of this adaptation is to force the centres to be as close as possible to as many vectors from the input space as possible. The value of α is initialised to a value of twenty, each time that the network is retrained, and its value is linearly decreased with the number of iterations until its value becomes 0; then the network is trained for a number of iterations (between 10 and 30 iterations for the whole training data set, depending on the time left for the training) in order to obtain the best possible weights for the final value of the centres.

A new centre is inserted into the network when the average error in the training data set does not fall more than 10% after 10 iterations (using the whole training set). In order to determine the most distant centre *C*, the Euclidean distance between each centre and each input vector is calculated and the centre whose distance from the input data vectors is largest is chosen. A new centre is inserted between *C* and the centre closest to it. Centres are also eliminated when they do not contribute significantly to the output of the neural network. Thus, a neuron is eliminated if the absolute value of the weight associated with that neuron is smaller than twenty per cent of the average value of the absolute value of the five smallest weights. The number of neurons in the middle layer is maintained above 20.

Results and Discussion

The approach presented in this paper combines the advantages of both connectionist and symbolic AI. The hybrid system has been tested in the Atlantic Ocean in September 1997 on a research cruise from the UK to the Falkland Islands. The cruise crossed several water masses and oceanographic fronts. The obtained results were very encouraging and indicate that the hybrid system is able to produce a forecast with an average error of 0.045 °C and with a probability of 94.4% that the error in the forecast is smaller than 0.1 °C. Only 5.6% of the forecasts have an error higher than 0.1 °C, 10.1% higher than 0.08 °C, and 27% higher than 0.05 °C

Although the experiment has been carried out with a limited data set (over a distance of 11000 km between the latitudes 50° North and 50° South), eleven water masses with different characteristics were traversed, six of them containing fronts; the Falkland Front, in particular, is one of the most chaotic oceanographic areas in the world. It is believed that these results are sufficiently significant to be extrapolated to the whole Atlantic Ocean.

The forecasting ability of the system is highest in areas with small instabilities and where there are many data profiles from which to choose in the retrieval stage of the CBR cycle. The forecast is less accurate in areas with large changes and many instabilities. The system is not able to forecast if there are no data profiles in the region where the forecast is made. In such a situation a time series ANN may produce a better result.

Experiments have also been carried out to evaluate the performance of the hybrid forecasting approach in comparison with several separate neural networks and statistical forecasting methods (Corchado and Lees, 1998): a Finite Impulse Response (FIR) model, an RBF network alone (trained with the data recorded during the 160 km previous to the forecast point), a linear regression model, an Auto-Regressive Integrated Moving Average (ARIMA) model and a CBR system alone (using the cases generated during the 160 km preceding the forecast point). Table 1 shows the average error in the forecast up to 5 km ahead using all of these methods.

Algorithm	*Type*	*Average Error*
FIR	ANN	0.091
RBF	ANN	0.103
Linear Regression	Statistics	0.131
ARIMA	Statistics	0.107
CBR	CBR	0.113
Hybrid CBR-ANN	CBR - ANN	0.017

Table 1 Comparison of methods (5km forecast)

The results of these experiments indicate that the forecasting errors outside the confidence limits are smaller with the hybrid CBR-ANN approach than with any of the other forecasting methods used. In particular, the forecasting error outside the confidence limits is less than 20% of the error produced by any of the other forecasting mechanisms.

The success of the system in generating effective forecasts may be attributed to the combination of an extensive database of past cases, supported by the neural adaptation mechanism which, each time around the forecasting cycle, enables the forecasting process to learn from all the selected closely matching cases.

The experimental results obtained to date are encouraging and indicate that the neural network supported approach is effective in the task of predicting the future oceanographic parameter values. Extrapolating beyond these results, it is believed that the approach may be applicable to the problem of parametric forecasting in other complex domains using sampled time series data.

Acknowledgement

The authors wish to express their gratitude to their colleagues Professor Jim Aiken and Mr Nigel Rees at Plymouth Marine Laboratory, for their support in this collaborative research.

References

Agre, G. and Koprinska, I. (1996). Case-based refinement of knowledge-based neural networks. In J. Albus, A. Meystel and R. Quintero (Eds.) *Proceedings of the International Conference on Intelligent Systems: A Semiotic Perspective* (Volume II).

Bishop C. R. (1995). *Neural Networks for Pattern Recognition*, Clarendon Press, Oxford.

Corchado J. M. (1995). Application of Filters, Masks and Artificial Neural Networks to Oceanographic Frontal Warning, BSc Honours Dissertation, University of Paisley.

Corchado, J. M., and Lees, B. (1998). A hybrid case-based model for forecasting, Department of Computing and Information Systems, University of Paisley.

Corchado, J., Lees, B., Fyfe, C. and N Rees, N. (1997). Adaptive agents: learning from the past and forecasting the future, *Procs PADD97 - First International Conference on the Practical Application of Knowledge Discovery and Data Mining*. London, 23rd to 25th April, 1997, Practical Application Co. Ltd., pp 109-123.

Garcia Lorenzo M.M. and Bello Perez R.E. (1996) Model and its different applications to case-based reasoning, *Knowledge-Based Systems*, Nov, Vol.9, No.7, pp.465-473 Elsevier Science B.V., Amsterdam, Netherlands.

Goodman, M. (1994). Results on controlling action with projective visualization. In *Proceedings of the Twelfth Naltional Conferenc on Artificial Intelligence*, pp. 1245-1250, Seattle, WA: AAAI Press.

Krovvidy, S. and Wee, W.G. (1993). Wastewater treatment systems from case-based reasoning, *Machine Learning*, Vol. 10, pp. 341-363.

Lees, B., Rees, N. and Aiken, J. (1992). Knowledge-based oceanographic data analysis, *Procs. Expersys-92*, editors: F Attia, A Flory, S Hashemi, G Gouarderes, J Marciano, IITT International Paris, October 1992, pp. 561-65.

Liu Z.Q. and Yan F. (1997). Fuzzy neural network in case-based diagnostic system *IEEE Transactions on Fuzzy Systems*, May 1997, Vol.5, No.2, pp.209-222.

Main J, Dillon T.S. and Khosla R. (1996). Use of fuzzy feature vectors and neural networks for case retrieval in case based systems, *Procs. Biennial Conference of the North American Fuzzy Information Processing Society - NAFIPS*, pp.438-443 IEEE, Piscataway, NJ, USA.

Reategui E.B., Campbell J.A. and Leao B.F. (1997) Combining a neural network with case-based reasoning in a diagnostic system *Artificial Intelligence in Medicine*, Jan, Vol.9, No.1, pp.5-27, Elsevier Science B.V., Amsterdam, Netherlands.

Rumelhart, D. E and McClelland, J. L. (1996). *Parallel Distributed Processing, Volume 1: Foundations*, MIT Press.

Quan Mao; Jing Qin; Xinfang Zhang; and Ji Zhou. (1994). Case prototype based design: philosophy and implementation *Procs Computers in Engineering 1994*. p. 369-74 vol.1, edited by Ishii, K. 11-14 Sept., Minneapolis, MN, USA, ASME New York, NY, USA.

Sun, R. (1996). Commonsense reasoning with rules, cases, and connectionist models: a paradigmatic comparison, *Fuzzy Sets and Systems*, Vol: 82 No: 2 pp. 187-200 Elsevier, Sept. 1996.

Appendix

1. Integration name/category:
real-time forecasting

2. Performance Task:
forecasting oceanographic parameter values

3. Integration Objective:
improved learning from multiple matching cases during case adaptation

4. Reasoning Components:
CBR;
radial basis function neural network

5. Control Architecture:
CBR as master;
neural network operation integrated into CBR cycle

6. CBR Cycle Step(s) Supported:
reuse (adaptation);
revision

7. Representations:
cases;
radial basis function weights and centres,
confidence limits

8. Additional Components:
DBMS

9. Integration Status:
empirically evaluated with real-time data;
results compared with alternative methods

10. Priority future work:
further evaluation;
application of the approach to other problem domains