Computing and Information Systems, Vol 5, No. 1, p.15-20

Neuro-Adaptation Method for a Case-Based Reasoning System

Juan Corchado, Brian Lees, Colin Fyfe, Nigel Rees and Jim Aiken

A Multi-agent approach is presented for identifying and forecasting the structure of the water ahead of an ongoing vessel. The work addresses the task of forecasting the behaviour of complex environments, in which the underling knowledge of the domain is not completely available, the rules governing the system are fuzzy and the available data sets are limited and incomplete. A hybrid approach is proposed that combines the ability of a Case-Based Reasoning System for selecting previous similar situations and the generalising ability of Artificial Neural Networks to guide the adaptation stage of the Case-Based Reasoning System. The successful application of the approach to oceanographic forecasting in the Atlantic Ocean is described.

1. INTRODUCTION

The World's Oceans have physical, biological and thermal characteristics that change seasonally and annually. An ocean's features change regularly; its location can vary several degrees in latitude or longitude (corresponding to a linear distance of 100km). The present knowledge of the structure of the oceans is at present too weak to create a full model of their behaviour; however, it is possible to model aspects of the behaviour of oceanic waters over limited areas. The "hybrid knowledge-based system" presented in this paper (Figure 1) has been designed to forecast the thermal structure of such water up to 40 km ahead of an ongoing vessel using a *Case-Based Reasoning* (CBR) System and an *Artificial Neural Network*. The Case-Based Reasoning System selects a number of cases (from a large case base) and the Artificial Neural Network retrains itself in real time in order to produce the final forecast.

The Artificial Neural Network (ANN) used in this investigation is a typical Radial Basis Function Network with characteristics that allow it to modify its internal structure depending on the characteristics of the waters in which the system is at any particular time. The learning method used in the system allows an *Autonomous Agent* to adapt to new situations and respond to oceanographic changes. The system produces more reliable forecasts than earlier methods. The paper outlines the oceanographic problem to be solved and discusses the CBR system and the Radial Basis Function ANN used for case adaptation. The forecasting strategy is compared with other methods and the application of the forecasting mechanism in the form of a multi-agent system is outlined.

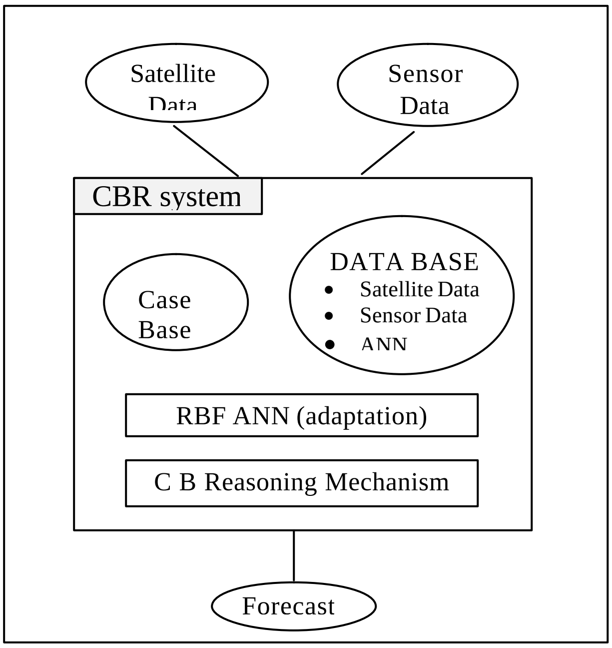

Figure 1: Hybrid knowledge system

2. THE OCEANOGRAPHIC ENVIRONMENT

The oceans are in a continuous state of flux. The scales of physical motion of the oceans and the atmosphere range from being ocean wide, through many intermediate sizes (hundreds or thousands of km) to finally, tiny eddies (in the range of 50 to 200km). At the edge of fronts (boundaries between water masses), small features between 10 to 50 km in amplitude can be detected. Between 1 and 10 km, local trends in the data set can be determined; movement of less than 1km represents the "noise" of the system. Furthermore, oceanic waters may be divided into largely homogeneous provinces, the variability of which can be easily described; Some provinces, e.g. the Arctic and Antarctic convergence zones, are extremely heterogeneous and are more variable; these provinces have their own general characteristics that can be described and simulated.

3. ARTIFICIAL INTELLIGENCE AND OCEANOGRAPHY SYSTEMS

The hybrid artificial intelligence (AI) system being investigated combines connectionist and symbolic techniques with the aim of successfully forecasting the water temperature ahead of an ongoing vessel. The hybrid system is implemented in the form of an autonomous agent operating on each ship; these agents can communicate with each other via a satellite system.

This hybrid AI approach to the problem of predicting the oceanic environment offers potential advantages over a conventional algorithmic data processing approach, as it is able to deal with uncertain, incomplete and even inconsistent data. With oceanographic data acquired in real time only, it is difficult to obtain an accurate forecast using either ANNs alone or conventional statistical techniques [1]; also, existing oceanic prediction models cannot easily represent small scale processes that are important in obtaining accurate forecasts in a particular oceanographic region. The limited data sets and associated uncertainties in oceanography are difficult to handle using conventional data processing approaches.

4. HYBRID SYSTEM

The aim of this work is to develop a forecasting tool able to produce accurate results anywhere, in any ocean, at any time. Figure 2 shows the relationship between the processes that are part of the CBR-hybrid system. The cyclic CBR process shown has been inspired by the ideas described by Aamondt and Plaza [2].

In Figure 2, shadowed words (together with the dotted arrows) represent the four steps of a typical CBR life cycle, the arrows together with the words in italics represent data flowing into, or out of the Case Base (situated in the centre of the diagram) and the text boxes represent the result obtained by each of the four stages of the CBR life-cycle. Solid lines show data flow and dotted lines indicate the order in which the processes that take part in the life cycle are executed.

Data are recorded in real time by sensors in the vessels and satellite pictures are received weekly. The *Knowledge Acquisition Module* is in charge of collecting, handling and *indexing* the data in the Case Base. Once the real-time system is activated on an ongoing vessel, a *new case* is generated every 2 km using the temperatures recorded by the vessel during the last 40km. This new case is used to **retrieve *m* cases** from a collection of previous cases.

The ***m retrieved cases*** are adapted by a neural network during the **reuse** phase to obtain an initial ***(proposed) forecast.*** Through the **revision** process the proposed solution is adjusted to generate the ***final forecast*** using the confidence limits from the knowledge base. ***Learning*** (**retaining**) is achieved by storing the proposed forecast and knowledge (ANN weights and centres) acquired by the ANN after the training and case adaptation.

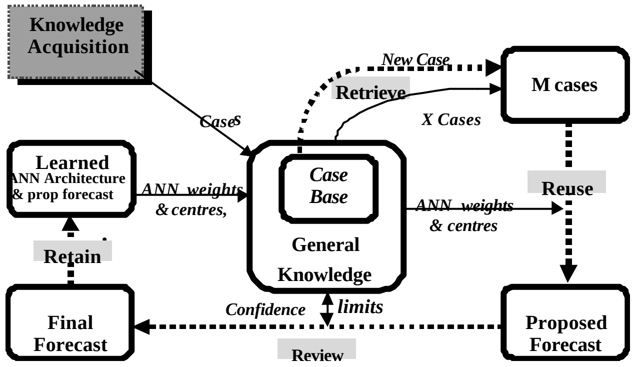

Figure 2: CBR-hybrid system architecture

4.1 Case Based Reasoning Operation

In Case-Based Reasoning systems [3] the solution to a problem is obtained by remembering a previous similar situation and by reusing information and knowledge relating to the solution of that previous problem. CBR has been successfully used in several domains, for example: diagnosis, prediction, control and planning [4]. The operation of CBR involves the adaptation of old solutions to match new experiences, using past cases to explain new situations, using previous experience to formulate new solutions, or reasoning from precedents to interpret a similar situation.

As explained in the previous section, ocean water features are very difficult to model and forecast using traditional techniques because of the lack of data, their incompleteness and the unpredictability of their changes. The models currently available can not accurately predict transient conditions.

4.2 Case Representation

Case representation involves specific knowledge about a particular situation. In applying CBR to oceanographic forecasting, a *case* is created to represent the 'shape' of a set of temperature values. There are two different types of cases:

<u>*Case Type A*</u> is composed of:

(i) a 40 km temperature input profile

Input profile: $x_0, x_1, .., x_k$, (where k = 40)

representing the structure of the water between the present position of the vessel and its position 40km back;

(ii) a 10km temperature output profile

Output profile: $y_0, y_1, .., y_q$, (where q = 10)

representing the structure of the water 10 km ahead of the present position of the vessel;

(iii) the latitude and longitude of the position of the vessel, time, day and year in which the data were recorded, and the tow orientation (North-South, South-East, etc).

Case Type B is composed of the same fields with the difference that the profiles are 160km

Input profile: $x_0, x_1, .., x_k$, *(where k = 160)*

and a 40km temperature profile

Output profile: $y_0, y_1, .., y_k$, *(where q = 40)*.

Cases of type A are used to forecast up to 10 km ahead and Cases of type B up to 40km ahead of the current position of the vessel. Cases are stored in a Case Base which is composed of millions of data profiles recorded by the Plymouth Marine Laboratory during the last decade, during many oceanic trips. Also, together with those cases, others have been created from satellite images in order to form a complete Case Base for the Atlantic Ocean. Profiles are also created from data downloaded by an *in situ* sensor working in real time.

4.3 Indexing Mechanism

The complexity and the quantity of the data that the system is capable of handling, requires a simple but rigid indexing mechanism in order to minimise the retrieval time of the cases.

Oceans have areas where the characteristics of the *water mass*, in terms of the evolution of the temperature, are relatively stable and the general characteristics of the water make them different from other water masses. Also the borders, limits or fronts between these water masses, can be located with relatively accuracy. Taking this into consideration, the indexing structure that is been implemented utilises the following data

Water Mass (Location and dimensions):

- **Trend associated to the Water Mass**
  Characteristics of the Trends
- **Satellite Images associated to the Water Mass**
  Characteristics of the S. Images

The water mass is either associated with a well known front or with a well defined area of the ocean. Each Water Mass has associated a number of data trends and satellite images, which name (if any), position (co-ordinates of the rectangle that contains the water mass or front), rules defining the cyclical change of location of the Water Mass (if known), etc.

By indexing the data in this way, it is possible to access in an easy way the relevant data needed at any particular point in time at a certain location. This indexing structure minimises the retrieval time.

4.4 Case Retrieval

Data are recorded in real time, and the *Input Profiles*, of both case types (40km and 160km), are created. A search, in the Case Base, is made for all the cases that are relevant within a radius of 3 degrees (300km) around the present situation of the vessel. Cases whose *Input Profiles* have the same orientation as the present one (i.e. North-South, North-West, South-East, etc.) are selected from this subset. These profiles are compared with the profile just recorded; three metrics are used by the retrieval algorithm with values that determine the similarity between the present case and each of the ones selected from the Case Base.

The three metrics used in the retrieval process were selected because each of them gives priority, in the retrieval, to cases based on different criteria that complement each other. These metrics (defined below) enable cases to be retrieved whose input profile is similar to the present one with respect to its general temperature similarity (*Gradient 1* and *Gradient 2*) and with respect to the general trend in temperature (*Gradient 3*).

- **Gradient 1**

This metric compares the shape of the present temperature profile with the shape of all the profiles stored in the case-base using the following equations.

G1a, the value of *Gradient 1* used to retrieve cases of type A, is given by

$$\sum_{i=0}^{7}\left(\left|\ \left(x_{39}-x_{(i*5)}\right)-\left(xa_{39}-xa_{(i*5)}\right)\ \right|*((84+(i*2))/100)\right)$$

where the vector *x* represents the present *Input Profile* for which a forecast should be retrieved, and *xa* represents each of the vectors retrieved by the CBR.

G1b, the value of *Gradient 1* used to retrieve cases of type B is given by

$$\sum_{i=0}^{15}\left(\left|\ \left(x_{159}-x_{(i*10)}\right)-\left(xa_{159}-xa_{(i*10)}\right)\ \right|*((70+(i*2))/100)\right)$$

where the vector *x* represents the present Input Profile for which a forecast should be retrieved, and *xa* represents each of the vectors retrieved by the CBR mechanism.

The gradients define the shape of each tow. The aim of this metric is to compare the structure of the present profile with the profiles stored in the Case Base. The smaller the value of the metric, the more similar will be the retrieved case to the *Present Input Profile*.

- **Gradient 2**

This metric is similar to the previous one, the difference being that the data is averaged using a window of 10% of the length of the *Input Profile.* Also, only the difference between the present temperature and one temperature value in every ten values, for cases of type A, (and one temperature value in every twenty values, for cases of type B) of each temperatures profile are used to calculate the value of the metric.

Gradient 2 gives a more general indication of the similarity between the present case and the retrieved ones than the previous metric.

- **Gradient 3**

The output of this metric is the absolute value of the difference between the gradient of the *Present Input Profile* and each of the cases retrieved from the Case Base. The gradient is calculated using the average value from the first value and the last 20% of the values of each *Input Profile*.

For a case of type A:

$$G3a = \left| \left(\left(\sum_{i=0}^{5} x_{i+35} - \sum_{i=0}^{5} x_i \right) / 5 \right) - \left(\left(\sum_{i=0}^{5} xa_{i+35} - \sum_{i=0}^{5} xa_i \right) / 5 \right) \right|$$

For a case of type B:

$$G3b = \left| \left(\left(\sum_{i=0}^{32} x_{i+127} - \sum_{i=0}^{32} x_i \right) / 32 \right) - \left(\left(\sum_{i=0}^{32} xa_{i+127} - \sum_{i=0}^{32} xa_i \right) / 32 \right) \right|$$

This metric shows the similarity in the general evolution of the water temperature between the present case and the ones retrieved from the Case Base.

5. Case Adaptation (reuse phase)

Work has been carried out elsewhere into hybrid systems in which CBR components co-operate with one or more reasoning elements [5]. In particular, there are a number of CBR based systems reported that use Constraint Satisfaction, Numeric Constraint Satisfaction, Model Based Reasoning, etc., for case adaptation.

Adaptation is one of the most difficult sections of the CBR cycle. Most adaptation techniques are based on generalisation and refinement heuristics. The novel approach, presented here, is based on ANNs and their ability to generalise. ***The ANN acts as a function that obtains the most representative solution from a number of cases which are the ones most similar to the current situation.***

The 30 best matches of each metric, presented in Section 4.4, are used to train a Radial Basis Function ANN in the adaptation stage. The algorithm developed for the construction of Radial Basis Function (RBF) networks used in this experiment is a variation of the general one [6]. In a Radial Basis Function ANN the input layer is a receptor for the input data. The hidden layer performs a non-linear transformation from the input space to the hidden layer space. 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*, which is the Euclidean distance between the inputs and the centres of the basis function, is calculated [7]. The weighted sum of the hidden neuron's activations is calculated at the single output neuron. The complexity of this ANN depends on the difficulty of determining which centres to use and where to locate them. The architecture presented in this paper automates this process and guarantees a number of centres very close to the minimum number that gives optimum performance.

The aim of the system is to forecast the water temperature up to 40km (i.e. at points 2.5, 5, 7.5, 10, 20, 30 and 40Km) ahead of a moving vessel. For this purpose two RBF ANNs are used: one uses cases of type A (to forecast up to 10km ahead) and the other uses cases of type B (to forecast between 10 to 40km ahead). Cases are coded in order to create the input and output vectors used to train the ANN.

The ANN trained with cases of type A (*ANN A*) uses nine input neurons, between 10 and 25 neurons in the hidden layer and one neuron in the output layer. The input data is a set of values, each of which is the difference between the last temperature (of the *Input Profile*) and the temperature values of the input profile taken every 4km. Only one neuron is used in the output layer to forecast up to 2.5km ahead; this output is fed back into the ANN in order to predict the following values up to 10km ahead. The output is the difference between the temperature at the present point and the temperature 2.5km ahead.

The ANN trained with cases of type B (*ANN B*) uses 15 neurons in the input layer, between 15 and 35 neurons in the hidden layer and one neuron in the output Layer. The input data is the gradient between the last temperature (of the *Input Profile*) and the temperature values taken every 10km ahead. Only one neuron is used in the output layer to forecast up to 10km ahead; this output is fed back into the ANN in order to predict the following values up to 40km. The output is the difference between the temperature at the present point and the temperature 10km ahead.

5.1 Initialisation

Initially, ten vectors are randomly chosen from the input data set and used as centres in the middle layer for the ANN of type A; fifteen vectors are chosen and are used for the ANN of type B. All the centres are associated with a Gaussian function. The width of the Gaussian for all the functions is set to the mean value of the Euclidean distance between the two centres that are further apart from each other.

5.2 Centre and Weight Adaptation

Training of the network is done by presenting pairs of input and desired output vectors. After an input vector activates every Gaussian unit to some degree these activations are propagated forward through the weighted connections to the output units which sum all incoming signals. The comparison of actual and desired outputs gives information for this input pattern from which the error is calculated using the Least Mean Square Rule [6].

The centre closest 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. An adaptation of this kind is particularly important because of the high-dimensional nature of the input layer. The value of α is initialised to 20 every time that the ANN is retrained, and its value is linearly decreased with the number of iterations until it becomes 0; then the ANN needs to be 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.

The *delta rule* [6] is used to adapt the weighted connections from the centres to the output neurons. In particular, for each presented pair of input and desired output vectors, one adaptation step is made, according to the delta rule.

5.3 Insertion of new units

A new centre is inserted in the network when the error computed by the ANN does not fall more than 10% after ten iterations. 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 in between *C* and the closest centre to it. Centres are also eliminated when they do not contribute significantly to the output of the ANN. Thus, if the absolute value of the weight associated with a neuron is smaller than 0.05, the neuron is eliminated. The number of neurons in the middle layer is controlled so as to be never less than 10 for *ANN A* and 15 for *ANN B*. This is a simple and efficient way of reducing the size of the ANN without decreasing its memory. This method is possible because the relevance of each weight increases as its absolute value increases.

5.4 Termination of training

ANN A is trained for 2.2 minutes and *ANN B* is trained for 3.3 minutes. In real time mode the ANNs need to produce a forecast every 2km (corresponding to 6 minutes for a speed of 12 knots, which is the maximum speed that the vessel can attain). After this time a new set of training cases is retrieved by the CBR and the ANNs are retrained. Therefore, even if the error is high the ANNs should produce a forecast. It has been shown, empirically, that these training times are sufficient to train the ANNs and to obtain an output average error of 0.04 with an error smaller than 0.08 in 98% of the situations.

The other two phases of the CBR cycle presented in Figure 2 are the *Review* and *Retain*. Phases. A review of cases is achieved by the use of confidence limits that modify the output of the ANN depending of the accuracy of the previous forecasts. During the retain phase the internal knowledge of the ANN (such as weights, centres, etc.) is stored in the data base.

6. The Autonomous System

Figure 3 shows the system architecture. Data is logged in real time; then it is both displayed and stored in a database. Every 2km the CBR mechanism retrieves

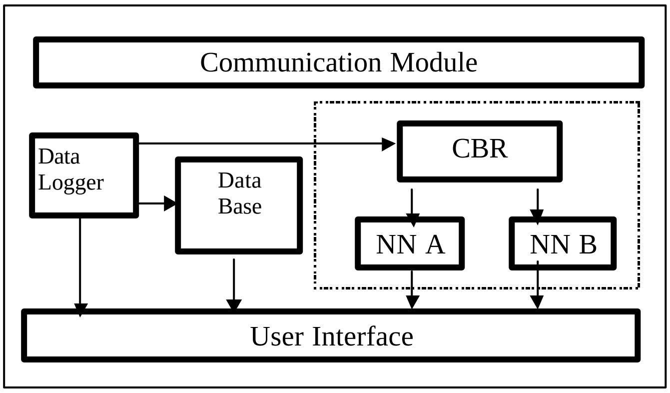

Figure 3: Autonomous agent architecture

the cases that match the most recent profile, based on the previously mentioned metrics, and passes them on to the ANNs, which produce the forecast. The Communication Module handles the communication between all the active agents of the multi-agent system.

7. Results and Conclusions

The methodology presented in this paper was tested using data recorded in the Atlantic Ocean in the summer of 1995. The obtained results were very encouraging and a single autonomous agent was built, installed in a vessel and tested in real time. The Case Base was loaded with satellite images recorded during the first week of September, 1997, and also with almost 2000 cruise tracks recorded during the last 3 years. The system was tested during the second week of September, 1997.

Distance ahead of the vessel (km)	*Average error in the ANN forecast.*	*Average Confidence Limit*	*Average. error after the review*
2.5	0.0098	0.011	0.0001
5	0.017	0.015	0.0011
7.5	0.025	0.023	0.0106
10	0.046	0.051	0.0124
20	0.156	0.170	0.0117
30	0.194	0.202	0.0132
40	0.285	0.246	0.0560

Table 1: Forecasting Results

Table 1 shows the average error in the prediction by the RBF ANN after the adaptation, together with the value of the confidence limits calculated during the review phase and the average error in the predictions outside the confidence limits. The results indicate the potential of the method and the advantage of using the ANN. To evaluate the accuracy of this method, these results have been compared with others obtained from a Finite Impulse Response (FIR) model [8], an RBF ANN (trained with the data recorded during the 160km previous to the forecast point), a linear regression model, an Auto-Regressive Integrated Moving Average (ARIMA) model and a CBR system (using the cases generated during the 160km previous to the forecast point). Table 2 shows the average error in the forecast up to 5km ahead using all of these methods.

Table 2 shows that, using the hybrid system, the error in forecasting up to 5km ahead is less than 20% of the error obtained using any of the other methods. For larger distances, the forecasting error with the hybrid method has been found to be between 15 to 50% of the error from any of the other methods. Its design as an autonomous system facilitates its implementation. A full multi-agent system is currently being developed with the aim of employing it on board several vessels.

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 2: Comparison of methods (5km forecast)

J. M. Corchado Rodriguez is a Researcher, and Dr. B. Lees and Dr. C. Fyfe are Senior Lecturers, at the University of Paisley. N. Rees and J. Aiken are at the Plymouth Marine Laboratory

References

[1] Lees B., Rees N. and Aiken J. (1992). *Knowledge Based Oceanographic Data Analysis*, Procs. Expert Systems 92, IITT Paris, pp 561-566

[2] Aamondt, A. and Plaza, E. (1994). *Case-Based Reasoning: foundational Issues, Methodological Variations, and System Approaches*, AICOM, Vol. 7 No. 1, March 1994.

[3] Kolodner J. (1993). *Case-Based Reasoning*, Morgan Kaufmann Publishers, Inc.

[4] López de Mántaras R. Plaza E., (1997). *Case-Based Reasoning: an overview*, AI Communications Vol. 10, pp. 21-29.

[5] Hunt J., Miler R., (1994). *Hybrid case-based reasoning*, The Knowledge Engineering Review, Vol. 9, No. 4, pp. 383-397.

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

[7] Fritzke B. (1994). Fast Learning with Incremental RBF Networks, Neural Processing Letters, Vol. 1, pp. 2-5.

[8] Corchado J., Fyfe C., Lees B. (1998). *Unsupervised Neural Method for Temperature Forecasting*, Proceedings of EIS, Tenerife, Spain.