CASE-BASED REASONING MODEL OF THE FISH DISEASE DIAGNOSIS

Zetian Fu [*] , Guoyong Hong [1] , Jian Sun [1]
[1] *Department of Mechanical Electrical Engineering, China JiLiang University,HangZhou, ZheJiang Province, P. R. China 310018*
[*] *Corresponding author, Address: China Agriculture University Key Laboratory of Modern Precison Agriculture System Integration Peking 100083*

Abstract: The system of Case Based Reason is researched in this paper, a kind of genetic algorithm is proposed to solve the problem of feature vector space, and proposes an approach of integration of CBR AND RBR, an example about fish disease diagnose inferring integration model is given.

Keywords: method, CBR , diagnose , reasoning

1. INTRODUCTION

In a general the knowledge of expert system have a direct impact on the performance of the traditional expert systems, because of the incomplete knowledge acquisition, to complete the construction and use of the reasoning rules for the system model is very difficult, construction expert knowledge acquisition system is the "bottleneck" problem. Since fish disease diagnosis system has the reasoning complexity and uncertainty, using case-based reasoning (CBR) approach can reduce the diagnosis in the process of man-made factors, the approach to solve the problem is very close to the human, the key is the method can memories the case, which is related to the most similar case when the system encountered a new problem, to a large extent the system is dependent on the quality of the case whether the reasoning ability of fish diseases can play its due role in this process, in order to get the better guarantee of objectivity and accuracy of the results, how to search the expected result in a becoming huge case is a key part of reasoning in retrieval CBR system.

2. THE BASIC CONCEPT OF CASE-BASED REASONING

2.1 The concept of CBR

CBR to develop artificial intelligence in the rule-based reasoning is different from a mode of reasoning, it refers to the use of the old cases or experience to solve the problem and to evaluate solutions to explain the abnormal situation or understand the new situation. Case is contextual information with the knowledge, the knowledge of the reasoning machine can achieve its objectives in the process, which can play a key role in the experience. In CBR, the current situation or the problems facing the case is known as the goal, but the memory of the case is known as the source of the problem or case.

2.2 The general process of CBR

CBR to deal with all the main elements are: identification of new issues, and finding a source of a similar case, the case is to use this new source of the problem a solution, the evaluation of this programme, and through case study and revise Library. CBR is actually a solution to the problem; learn from experience, to resolve the issue of a new cycle and integration process. General CBR cycle process includes the following four: case retrieval, reuse case, case modification and case studies. Case retrieval is retrieved from the case with most new issues similar to the case; reuse is reusable case to case retrieval of information and knowledge to solve new problems; case for change is modification programme; case study is to learn new Experience and put it into the existing case library. Figure 1 can be used to describe a cycle of the CBR.

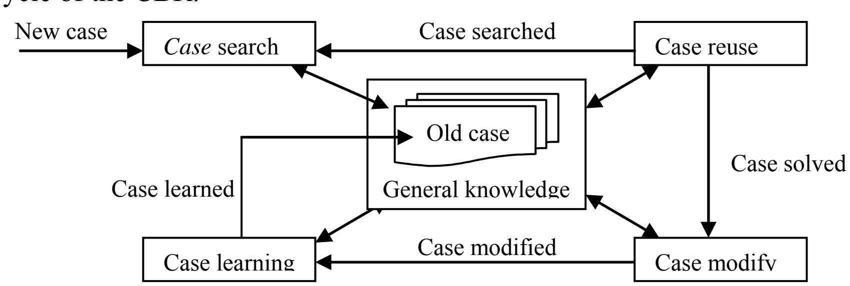

Fig.1: CBR Cycle

According to the above description, CBR can be found on the core issue includes five aspects: that case method, case retrieval method, case reuse methods, case modification method and case study methods.

3. THE CONTENT OF FISH DISEASE DIAGNOSIS

Fish Disease Diagnosis is through a variety of ways and means of observing and monitoring the fish's living conditions and in accordance with the known symptoms, were diagnosed to identify possible targets of fish or to be a disease, and led to the direct cause of these diseases, to ensure that Was diagnosed target fish in certain seasons and the living environment of healthy growth. Fish Disease inference system refers to the completion of a diagnostic function of the fish's computer system, and its main task is under observation by the user to show the fish and its related symptoms, such as the quality of information, identifying the fish living in water may exist disease and its related properties, and to determine the corresponding control measures or strategies.

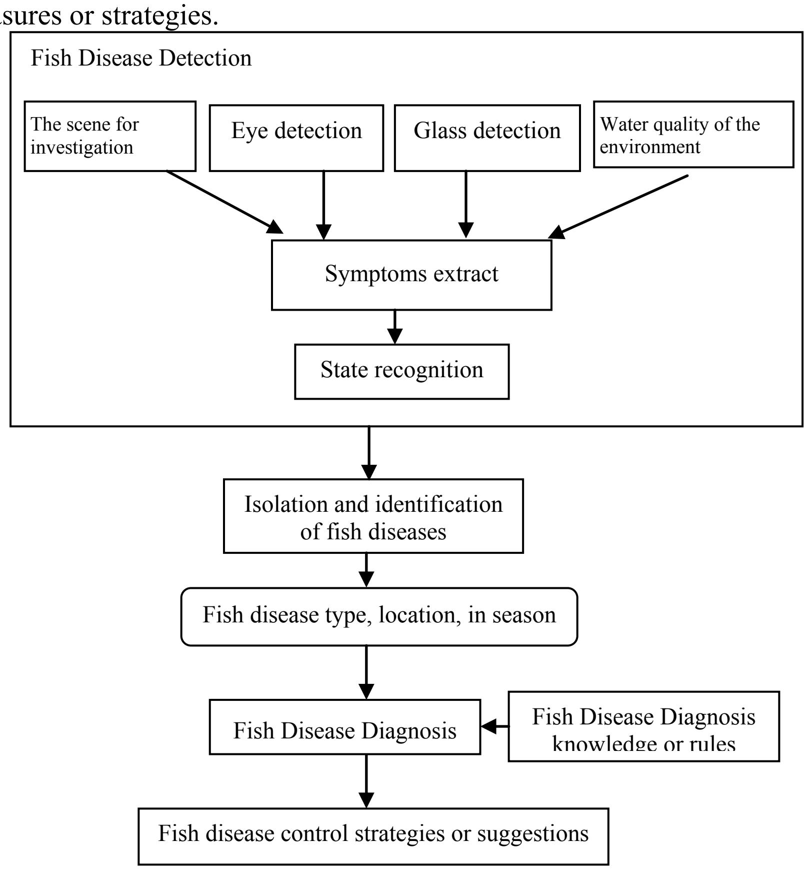

Fig.2: Fish Disease Diagnosis process

Fish Disease Diagnosis has three basic attributes, that is, symptoms, diseases and causes. One is the external manifestations of the disease or the results of the symptoms of the disease occurred before or during the course of fish in certain areas may show the abnormal phenomenon, which is the cause of diseases leading to the fundamental attributes collection. Therefore, a fish disease diagnosis system in the specific diagnosis of the process is shown in Figure 2.

4. CASE EXPRESSION OF FISH DISEASE DIAGNOSIS REASONING

Normally, a typical case should at least include a description of the problem and a description of the solution. For fish disease diagnosis system, the problem is demonstrated by the fish's symptoms, such as water quality parameters; solution is to identify the disease, causes, and gives the corresponding treatment.

In the fish disease diagnostic reasoning system, according to the characteristics of fish diseases and content to eigenvector as a case of that form, in order to facilitate the narrative, we head to the seizure process as an example, omitted the specific case had little to do with the case-based reasoning Phenomenon and disease control methods, such as content, only concerned with fish diseases related to the case of attributes.

We use muscle, skin, abdomen, scales, head, fin, gill Department, intestinal and other abnormal symptoms to create a disease case. Listed in Table 1 is the basis of these characteristics established by the Lei Yue Mun in the five disease cases, in each case on behalf of its state of the disease demonstrated by the different characteristics of the portfolio. In these cases did not consider the weight characteristics of the impact of the case, that is the assumption that all cases of the role and influence in the same.

Table1 Carp disease case

ID	Disease	Muscle	Surface	Abdominal	Scales	Head	Fin	Gill	Intestinal
Case1	Baekdu-mouth	000	A01	000	000	B04	000	000	000
Case2	Enteritis	000	000	D01	000	000	000	000	H03
Case3	Red skin	000	A04	000	E01	000	F01	C03	000
Case4	Black gill	000	000	000	000	B01	000	C01	000
Case5	Scabies	000	A07	000	000	000	F02	000	000

If each case as a row vector, from the above five cases of the case through the matrix that can be set as follows:

$$C = \begin{bmatrix} 000 & A01 & 000 & 000 & B04 & 000 & 000 & 000 \\ 000 & 000 & D01 & 000 & 000 & 000 & 000 & H03 \\ 000 & A04 & 000 & E01 & 000 & F01 & C03 & 000 \\ 000 & 000 & 000 & 000 & B01 & 000 & C01 & 000 \\ 000 & A07 & 000 & 000 & 000 & F02 & 000 & 000 \end{bmatrix} \quad (1)$$

In practice, the different characteristics of the role and impact of cases there is a certain degree of difference, but this difference in performance under different circumstances to different degrees and others. Table 1 listed in the case cited, if the consideration of the characteristics of eight cases of the impact and role of the differences, they give different weights, and you can get a new set of cases, as shown in table 2.

Table 2 Carp disease case

ID	Disease	Muscle	Surface	Abdominal	Scales	Head	Fin	Gill	Intestinal
Case1	Baekdu-mouth	000	A01	000	000	B04	000	000	000
	Weight	0	0.5	0	0	0.5	0	0	0
Case2	Enteritis	000	000	D01	000	000	000	000	H03
	Weight	0	0	0	0	0	0	0	0.5
Case3	Red skin	000	A04	000	E01	000	F01	C03	000
	Weight	0	0.25	0	0.25	0	0	0.25	0
Case4	Black gill	000	000	000	000	B01	000	C01	000
	Weight	0	0	0	0	0.5	0	0.5	0
Case5	Scabies	000	A07	000	000	000	F02	000	000
	Weight	0	0.5	0	0	0	0.5	0	0

As in the case of different features and characteristics of the weight may change, as in the example of Black gill disease symptoms and weight in the performance of the head and gills, and enteritis disease symptoms and weight are reflected in the abdomen and bowel, It is an extraordinary collection of configuration case, we use this case the largest collection of feature set to solve the unusual configuration problems.

We can organize the collection features the largest case characteristic matrix:

$$F = \begin{bmatrix} f_{1,1} & f_{1,2} & \cdots & f_{1,m} \\ f_{2,1} & f_{2,2} & \cdots & f_{2,m} \\ . & . & . & . \\ f_{n,1} & f_{n,2} & \cdots & f_{n,m} \end{bmatrix} \quad (2)$$

In characteristic matrix, each line represents a case, each column represents a case of different characteristics in case the characteristics of the matrix of every element of the value $f_{i,j}\,(1 \le i \le n, 1 \le j \le m)$, and we do the following:

$$f_{i,j} = \begin{cases} \text{the value of characteri stic j in case i, when characteri stic j} \in F \\ \text{no definition , when characteri stic j} \notin F \end{cases}$$

When the feature is a feature of the case, it said in the case of the characteristics of value, or no practical significance, can be used arbitrarily to fill the appropriate value of the element in the matrix in the position.

Similarly, we have the characteristics of the case in the weight also similar manner in accordance with the above organizations to become a matrix form, saying the matrix for the characteristics of the right matrix.

$$R = \begin{bmatrix} r_{1,1} & r_{1,2} & \cdots & r_{1,m} \\ r_{2,1} & r_{2,2} & \cdots & r_{2,m} \\ . & . & . & . \\ r_{n,1} & r_{n,2} & \cdots & r_{n,m} \end{bmatrix} \quad (3)$$

Where: R is the characteristics of the right matrix; n is the collection C features the largest number of features, the case for the number of cases in the collection.

Characteristics of the matrix R, each row represents a case of different characteristics in the case of the characteristics of the right values, the matrix R of every element of the value is $r_{i,j} (1 \le i \le n, 1 \le j \le m)$, we do the following:

$$f_{i,j} = \begin{cases} \text{the value of characteri stic j in case i, when characteri stic j} \in F \\ \text{no definition , when characteri stic j} \notin F \end{cases}$$

When the feature j is a feature of the case i, said $r_{i,j}$ is the right value of characteristics, otherwise $r_{i,j} = 0$.

In the above definition on the basis of a case that could set the case matrix collection C features and characteristics of the matrix right to buy the product matrix:

$$C = F \bullet R^T \quad (4)$$

For example, we can use matrix shown in table 2 from the collection of fish diseases of the case said:

$$\begin{bmatrix} C1 \\ C2 \\ C3 \\ C4 \\ C5 \end{bmatrix} = \begin{bmatrix} 000 & A01 & 000 & 000 & B04 & 000 & 000 & 000 \\ 000 & 000 & D01 & 000 & 000 & 000 & 000 & H03 \\ 000 & A04 & 000 & E01 & 000 & F01 & C03 & 000 \\ 000 & 000 & 000 & 000 & B01 & 000 & C01 & 000 \\ 000 & A07 & 000 & 000 & 000 & F02 & 000 & 000 \end{bmatrix} \bullet \begin{bmatrix} 0 & 0.5 & 0 & 0 & 0.5 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.5 \\ 0 & .25 & 0 & .25 & 0 & 0 & 0.5 & 0 \\ 0 & 0 & 0 & 0 & 0.5 & 0 & 0.5 & 0 \\ 0 & 0.5 & 0 & 0 & 0 & 0.5 & 0 & 0 \end{bmatrix} \quad (5)$$

5. FISH DISEASE DIAGNOSIS REASONING OF THE CASE RETRIEVAL SYSTEM

In general, CBR system is suitable for experienced, but the lack of knowledge of the field, while RBR system can handle rich and knowledge of the system on the field. Therefore, the CBR and RBR will be integrated into one of the two systems will have the strengths to overcome the shortage of both, so that the entire system to achieve a higher level of intelligence, is of great significance. It will not only greatly enhance the flexibility and comprehensive reasoning ability, but also significantly reduce the case retrieval and the burden of the case.

We will go into two tiers: the bottom of all cases; upper deck is a typical case, each case represents the typical one or several months and it is very similar to the case, typical case of the amended rules is essential, it should be the guidance of experts in the field. To be good to amend the rules, for a similar case semantic difference between the measures is essential, semantic word table can be defined by the index of keywords to describe, indexing vocabulary of keywords can be used in case the same function and the description of semantic network nodes. Application of amended rules also requires case assessment that the subsystem is the evaluation of the answer, because once the amendment is not likely to be the result of the establishment, but also further changes. Sometimes, the problem can not be with a certain rules of the first match, now need to separately image to the first of several rules, the final rules will be several interpretations of the results integrated into the whole question to answer.

5.1 Case index

The index case mechanism is an important issue, the objective is to provide a case for the search mechanism, to make the retrieval in the future in line with the need to quickly identify the cases or cases set. The index should be specific, clear and easy identification. Case of the index can have multiple, respectively, for different purposes. Search algorithms should be able to use the index case in time to meet the constraints to identify the premise of the case. This means that the case should be indexed; making the case in any need at all times can be found.

In this paper, C-means clustering algorithm to a polymerization center on behalf of a disease, according to the type of disease the number is divided into several polymerizations centre, the centre for polymerization index.

C means clustering model, established the objective function:

$$J = \min \sum_{r=1}^{P} \sum_{i=1}^{m_r} \| X_j^{(r)} - C_r \|^2 \tag{6}$$

Where: Cluster Center

$$C_r = \frac{1}{m_r}\sum_{i=1}^{m_r} X_i^{(r)} \quad (i = 1,2,..., \ r = 1,2,...P) \ , \ \sum_{r=1}^{P} m_r = N \qquad (7)$$

where: m_r is the number of r sample； $X_i^{(r)}$ express sample X_i belong to r； N is the number of sample； P is the number of cluster center(2≤P≤N-1).

We use genetic algorithms to solve the problem dynamic clustering.

Step 1 chromosome structure

Genetic algorithms set up the relevant parameters, max_gen: the largest number of iterative; N: population size; length: the length of chromosomes; Pc: cross-probability; Pm: mutation probability; P: initial cluster centres; α: the fitness function of Parameters;

Step 2: groups initialization
for i = 1 to N do
for j = 1 to length do
Xi the first-chromosome gene = random (0, P);
Endfor
Endfor

STEP 3: Calculation of fitness function
for i = 1 to N do
Xi chromosome calculation of the objective function Ji;
endfor
According to Xi chromosome target function of the value of Ji sort;
for i = 1 to N do
Individual fitness calculation
endfor
for j = 1 to N do
An Introduction to calculate the cumulative
endfor

Step 4: The Executive select Options

Step 5: the implementation of cross-operation

Step 6: The Executive mutation

Step 7: retain the elite strategy
if (Minimum of NewGeneration <Minimum of OldGeneration) then
Father and with the best of the individual offspring to replace the worst of the individual;
endfor

Step 8: The dynamic changes in the number of cluster
If (the best group of individuals in the M-generation, has not changed) then
P = Popt + n, and to Step 2;
else
Running;
endif

Step 9: termination of the conditions tested
if (gen <max_gen) then
Order gen = gen +1, and turn to Step 3;
else
Stop, the output results
endif

5.2 Rule-based and case retrieval algorithms

Rule-based and case search algorithm is as follows:

1: input new case index generated in accordance with the relevant rules, create a new case description of the rules;

2: judge rules coding RULE_CODE the degree of similarity, if the same case were directly targeted. If not identical, the value of similar properties in the area threshold if the threshold value, then retrieves the corresponding case set; if the threshold, based on the similarity case match and then retrieves the corresponding The case. If the search to retrieve cases of unreasonable or no case, to 4;

3: In Search of the process, the preservation of relevant case study of key knowledge and interest in hobbies, credited to the inspiring knowledge base.

4: selected cases involving the target system (object to sort through), in accordance with our neighbors to improve the method of calculating targets under the similarity measure;

5: Does also involves other object, if so, the system for cases involving the next target, or else to 5;

6: calculate the case of the similarity system;

7: The similarity of the cases is the highest result;

8: The Case users or experts satisfied with.

If satisfied, such cases will be listed as a candidate case set, the end of reasoning.

6. CONCLUSION

The system based on the characteristics of fish diseases and content of case-based reasoning is studied in this paper, we can effectively solve the problems in the diagnosis of fish diseases, and improve the operating efficiency of the system. CBR has a crucial impact on the accuracy and practical application, but it is difficult to determine the right value, based on random genetic algorithm global search capability, a genetic algorithm to complete the space characteristics of the right to direct the search algorithm is presented, For people to solve practical problems is not simply relying on experience, not simply relying on theoretical knowledge, is often a combination of both the facts, it is a combination of methods, and practice shows the method is effective.

ACKNOWLEDGEMENTS

Funding for this research was provided by China Agriculture University Key Laboratory of Modern Precision Agriculture System Integration (P. R. China). The first author is grateful to China Agriculture University for providing her with pursuing a PhD degree.

REFERENCES

C. R. Sukumaran, B. P. N. Singh. Compression of bed of rapeseeds: the oil-point, Journal of Agricultural Engineering Research, 1989,42:77-84

Cheng Wang. Yinghong LI. Hengxi Zhang. The improvement of CBR technology in the diagnosis of failure. Computer Engineering and Applications. 2003（15）：44～46

E. Davion, A. G. Meiering, F.J. Middendof. A theoretical stress model of rapeseed, Canadian Agricultural Engineering, 1979, 21(1): 45-46

J. B. Noh. A case-based reasoning approach to cognitive map-driven tacit knowledge management. Expert Systems with Application. 2000(19)：249～259

J. M. Garrell. I. Guiu. Automatic diagnosis with genetic algorithms and case-based reasoning. Artificial Intelligence in Engineering. 13(1999)367～372

Pi Sheng Deng. Using case-based reasoning approach to the support of ill-structured decisions. European Journal of Operational Research. 1996 (93)511～521

S L.A Salzburg, Nearest Hyper rectangle Learning Method. Machine Learning. 1991,6:251-276.

S M.A Bryant Case-based Reasoning Approach to Bankruptcy Prediction Modeling

T. W. Liao. A case-based reasoning system for identifying failure mechanisms. Engineering Applications of Artificial Intelligence. 2000 (13)：199～213