

A CBR SYSTEM FOR FISH DISEASE DIAGNOSIS

Wei Zhu, Daoliang Li^{*}

Key Laboratory of Modern Precision Agriculture System Integration Research, China Agricultural University, 17 Tsinghua East Road, Beijing, China

^{} Corresponding author, Address: P.O. Box 121, EU-China Center for Information & Communication Technologies, China Agricultural University, 17 Tsinghua East Road, Beijing, 100083, P. R. China, Tel: +86-10-72736717, Fax: +86-10-62737679, Email: li_daoliang@yaohoo.com*

Abstract: Fish disease diagnosis is a complicated process and requires high level of expertise. However, there's no accepted general knowledge in fish disease diagnosis. This paper describes a CBR (case-based reasoning) system for fish disease diagnosis. A two-step case retrieve model is proposed in this paper.

Keywords: case-based reasoning, fish disease diagnosis, nearest neighbor

1. INTRODUCTION

Fish disease diagnosis is a complicated process and requires high level of expertise. This problem domain has its own problems, the major one being that the effort to deal with a multitude of diseases for multiple species, and another one being that there's no accepted general knowledge in fish disease diagnosis (Daniel Zeldis, 2000).

An expert system for fish disease diagnosis called Fish-Expert has been developed. Fish-Expert, in which rule-based reasoning is applied, can mimic human fish disease experts (Daoliang Li, 2002). There are two deficiencies with the Fish-Expert, the major one being that rule-based reasoning requires some 'deep' knowledge in order to be truly effective, however there's no accepted general knowledge, another one being that inference engine is too complex to work efficiently and is time-consuming.

This paper describes a CBR system for fish disease diagnosis. CBR is a well recognized and established method for building medical expert systems [4]. Instead of relying on general domain knowledge, CBR uses the storage of a large number of previously solved cases (Isabelle Bichindaritz, 2006). Through the interview with these fish experts, plenty of fish disease cases those can be used in CBR have been acquired. And a two-step case retrieve model is proposed. Some experiences with developing CBR system are discussed and the conclusions are provided at the end of the paper.

2. SYSTEM ARCHITECTURE

The overall architecture of the system is show in Figure 1. Through a forms interface, the fish farmers' requirements could be input. And the CBR system will search the fish disease case base of past cases and retrieve similar cases. Details of the similar case will then be available to the fish farmers. All this information would then be automatically passed back to an agent to authorize or change, if necessary. Once a case is completed, its details would be added to the fish disease case base (Isabelle Bichindaritz, 2006).

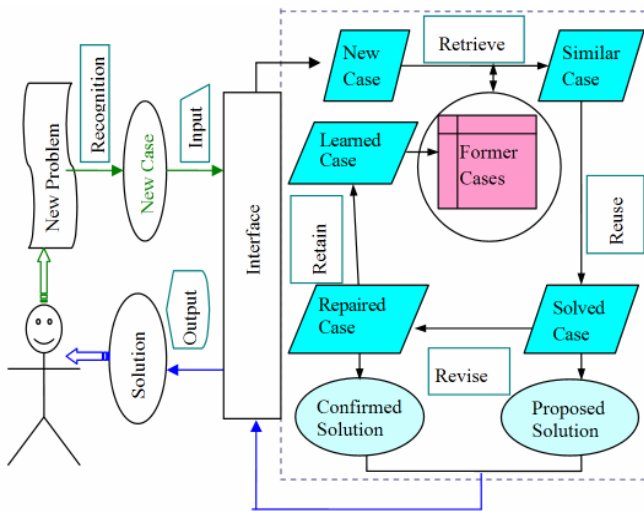


Figure 1. System architecture

3. METHODOLOGY

In the following, how the CBR to be applied in fish disease diagnosis is described. The subtasks are referred to as case representation, case indexing, case retrieval and case reuse.

3.1 Case Representation

A case of fish disease diagnosis is a set of empirical data, gathering the experience of the fish vets involved in a previous situation. The fish vets write the usual description of a case in natural language in a general standardized report. From the analysis of a corpus of available reports, a model has been developed. Within this model, a case of fish disease diagnosis is consisted of (Object, Symptoms, Treatments). As there is no fish disease vocabulary standard available for descriptions, the domain-related items have been chosen by consensus between some fish vets (Anil Varma, 1999). The symptoms are described as (0-1) vector, where 0 means the symptom doesn't appear, and 1 means the symptom appears.

3.2 Case Indexing

A two-step case indexing model is proposed to quickly locate similar previous cases in this paper. In the first step index, a clustering algorithm is used to partition all past cases in the fish disease case base into clusters according to their design specifications. Primary index features are part of the explanation of the new case input by the fish farmers. It initially identifies which case base the new case belongs to. With the indexing mechanism conducted in advance, the case retrieval procedure can be accelerated. In the second step, the observed features themselves are used as secondary features only. The similarity between the new case and the small set of retrieved cases of the first step is calculated using this simple nearest neighbor algorithm (Florian Hartge, 2006).

3.3 Case Retrieval

The retrieve task starts with a partial fish disease symptoms description, and ends when a best matching previous fish disease case has been found. Its subtasks are referred to as Identify Features, Initially Match and Select, executed in that order (Isabelle Bichindaritz, 2006).

3.4 Identify Feature

To identify a new fish disease case whose input symptoms features should be noticed. Unknown symptoms features may be disregarded or requested to be explained by the fish farmer. To understand a new case involves filtering out noisy symptoms descriptors, inferring other relevant problem features, checking whether the feature values make sense within the context, etc (Kyung-Sup Kim, 2001).

3.4.1 Initially Match

Once the fish disease symptoms features have been identified, the Initially Match process starts. Finding a set of matching disease is done by using the symptoms features as indices to the fish disease knowledge base in a direct way. The possible fish disease set is retrieved solely from input symptoms features. A way to assess the degree of similarity is needed, and several similarity measurements have been proposed, based on surface similarities of fish disease symptoms features (Nirmalie Wiratunga, 2004).

3.4.2 Select

The selection process typically generates consequences and expectations from each retrieved case, and attempts to evaluate consequences and justify expectations. The similarity between the new case and the small set of retrieved cases is calculated using this simple nearest neighbor algorithms. From the set of similar cases, a best match case is to be chosen. The best matching case is usually determined by evaluating the similarity of the fish disease symptoms features between the new case and the small set of retrieved cases. The case that has the strongest explanation for being similar to the new problem is chosen (Isabelle Bichindaritz, 2006). In this stage, the small set of retrieved fish disease cases is compared by the client-side applet with the original query and similarity is calculated using this simple nearest neighbor algorithm (Abdus Salam Khan, 2003).

3.5 Case Reuse

The reuse of the retrieved fish disease case solution in the context of the new case focuses on two aspects: (a) the differences among the previous and the new case and (b) what part of a retrieved case can be transferred to the new case (I. Watson, 1999). The modification is done by using the fish disease diagnosis knowledge in this paper.

4. CONCLUSION

In conclusion, CBR is applied as the inference engine to mimic human fish experts. Instead of relying on general fish disease knowledge which is lacking, CBR uses the storage of a large number of previously solved cases. And a two-step case indexing model is proposed to quickly locate similar previous cases. Based on the two-step case indexing model, an efficient case retrieval procedure is developed to find similar cases from the fish disease case base for a new case. Experience has shown that CBR approach has been able to contribute significantly in fish disease diagnosis. Though the system is developed specific for fish disease diagnosis, it can also easily be developed available for other domains.

ACKNOWLEDGEMENTS

The authors would like to thank many fish experts from Beijing Aquaculture Science Institute and Aquaculture department of Tianjin Agricultural College for their fruitful comments and suggestions. The authors acknowledge the Fok Ying Tung Education Foundation (Grant No. 94032) for funding our research.

REFERENCES

- Abdus Salam Khan, Achim Hoffmann. Building a case-based diet recommendation system without a knowledge engineer. *Artificial Intelligence in Medicine*, 2003(27): 155-179
- Anil Varma, Nicholas Roddy. ICARUS: design and deployment of a case-based reasoning system for locomotive diagnostics. *Engineering Applications of Artificial Intelligence*, 1999(12): 681-690
- Daniel Zeldis, Shawn Prescott. Fish disease diagnosis program-problems and some solutions. *Aquacultural Engineering*, 2000(23): 3-1
- Daoliang Li, Zetian Fu, Yanqing Duan. Fish-Expert: a web-based expert system for fish disease diagnosis. *Expert Systems with Applications*, 2002(23): 311-320
- Florian Hartge, Thomas Wetter, Walter E. Haefeli. A similarity measure for case based reasoning modeling with temporal abstraction based on cross-correlation. *Computer Methods and Programs in Biomedicine*, 2006(8): 41-48
- I. Watson. Case-based reasoning is a methodology not a technology. *Knowledge-Based Systems*, 1999(12): 303-308
- Isabelle Bichindaritz, Cindy Marling. Case-based reasoning in the health sciences: What's next? *Artificial Intelligence in Medicine*, 2006(36): 127-135
- Isabelle Bichindaritz. Memoire: A framework for semantic interoperability of case-based reasoning systems in biology and medicine. *Artificial Intelligence in Medicine*, 2006(36): 177-192
- Kyung-Sup Kim, Ingoo Han. The cluster-indexing method for case-based reasoning using self-organizing maps and learning vector quantization for bond rating cases. *Expert Systems with Applications*, 2001(21): 147-156
- Nirmalie Wiratunga, Susan Craw, Bruce Taylor, Genevieve Davis. Case-based reasoning for matching Smarthouse technology to people's needs. *Knowledge-Based Systems*, 2004(17): 139-146