# Measures of Solution Accuracy in Case-Based Reasoning Systems

William Cheetham and Joseph Price

General Electric Company,
1 Research Circle, Niskayuna, NY 12309
{cheetham,pricejo}@research.ge.com

**Abstract.** The case-based reasoning (CBR) methodology can be augmented with the ability to determine the confidence in the correctness of individual solutions. A confidence calculation can be added to the REUSE portion of the CBR methodology. The confidence calculation takes confidence indicators, like "number of cases retrieved with best solution" and "average similarity of cases which suggest an alternative solution," and generates a confidence value. The information gain algorithm C4.5 can be used to select the best confidence indicators by evaluating their usefulness in historical cases. A genetic algorithm can be used to optimize and maintain the confidence calculation.

## 1 Introduction

Mark Twain is quoted as saying, "I was gratified to be able to answer promptly, and I did. I said I didn't know." If a person is asked a question and they do not know the answer, they are usually able to say that they are not confident in their ability to give a correct answer. When a person does not have confidence, they may not be able to answer at all (by saying "I don't know") or they may supply multiple guesses at what the answer could be. The ability to know the limits of one's knowledge is valuable in an intelligent system. This paper describes how a Case-Based Reasoning (CBR) [1, 9] system can be constructed to provide a measure of its solution accuracy that says when it is confident in its answer, when it is not able to answer the question, or when there are multiple possible answers.

Our primary motivation for creating an intelligent system that can determine its confidence in each solution it produces is to improve efficiency of business processes by automating actions when a CBR system has high confidence in its solution. Knowledge-based decision tasks that can benefit from automation include financial tasks such as granting or rejecting applications for insurance or credit and service tasks such as diagnosing or predicting failures with machinery. All of these tasks are currently performed by experts and are difficult to automate because some instances of each task can be quite complicated. An intelligent system that knows when it should and should not be used to automate a task can be very valuable because it can have a low error rate on the tasks that it does automate and pass the tasks where it is more likely to be incorrect on to the person who had been performing that task. This would free

P. Funk and P.A. González Calero (Eds.): ECCBR 2004, LNAI 3155, pp. 106–118, 2004.

the person from the burden of dealing with routine instances of a task and allow them to focus on the unique and difficult to automate instances.

If there were always a person evaluating the output of a CBR system and deciding what action to perform, then similarity values, which most CBR systems output, would probably be sufficient for the person to make their decision. However, when the person is removed from the loop, more certainty that a specific action should be performed then is available from just the similarity values is often required. Having the CBR system determine the predicted accuracy of its solution is one way to provide the certainty that is needed to automate the task. There can be large financial benefits from automating even a percentage of the occurrences of expensive decision tasks. Other benefits of the automated process include speed, consistency, and the ability to monitor and optimize the process. These benefits, financial and otherwise, have driven General Electric to automate multiple decision tasks by using a CBR system with an integrated confidence calculation. Some example applications are described in section 5.

There are many other reasons that it would be beneficial to have intelligent systems that can give "I don't know" as a possible answer. First, if a system is presented with a problem that it was not designed to solve and is forced to give an answer it would probably give an inaccurate one. Second, the domain for which the system was created might change over time. If the system was forced to give an answer after the domain changed, there could be a situation when the system would not be able to give an informed answer, but will still give an answer. Third, if there is an alternative method of determining an answer, such as using a different system or having a person create it manually, then it would be good to know when these alternatives should be used instead of the usual system.

Before determining how a CBR system can produce a measure of confidence in its output, we should first observe how humans determine confidence in their decisions. Specifically, since we are determining confidence for CBR systems, we should analyze how humans determine confidence using experience. If a person has no experience relating to a decision they would have low confidence in any guess at a solution. If they had many previous experiences with similar problems, and if all of these experiences indicated the same solution could be applied to this problem, then they would have high confidence in that solution. If a person has conflicting evidence regarding the solution to a problem then their confidence level would be reduced to some 'medium' level. One method they might use to determine confidence is to list the evidence supporting the solution and the evidence against the solution, then determine if the supporting evidence outweighs the contrary evidence. There is also the possibility of giving multiple answers when there is no clear winner and there are multiple potential solutions. It is bad for people to have either too little of too much confidence in the decisions. Someone who has too little confidence can be thought of as timid or unsure of himself. Someone who has too much confidence can be thought of as arrogant or conceited.

Different problems could require different levels of certainty in order to have high confidence. If there is little harm done by giving an incorrect answer then a low level of certainty would be needed to provide an answer. An example of this type of CBR