# SIMILARITY LOCAL ADJUSTMENT: INTRODUCING ATTRIBUTE RISK INTO THE CASE

**Juan L. Castro,** Department of Computer Science and Artificial Intelligence, ETSI Informática ,Granada University, castro@decsai.ugr.es

**María Navarro,** Department of Computer Science and Artificial Intelligence, ETSI Informática, Granada University, minj@correo.ugr.es

**José M. Sánchez,** Department of Computer Science and Artificial Intelligence, ETSI Informática, Granada University, jmsa@decsai.ugr.es

**José M. Zurita,** Department of Computer Science and Artificial Intelligence, ETSI Informática, Granada University, zurita@decsai.ugr.es

**Abstract**

*The method used to evaluate the overall similarity between cases in the CBR retrieval stage plays a very important role when deciding the case to select and therefore the final solution to apply. Many case retrieval techniques have been developed and perhaps the most popular is the one which uses the nearest-neighbour (NN) matching function. This technique first calculates the similarity between the target problem and an old case regarding individual attributes, and then places the overall similarity of the target problem with the old case. The overall similarity is assessed by a weighted sum of all the similarity measures between the attributes, where the weighting factor represents the degree of importance of each attribute. Since this factor is sometimes allocated by a human expert and other times by the human user, the weight assessments usually involve human subjectivity. In order to prevent this subjectivity from affecting case retrieval and to avoid losing the valuable information obtained when measuring the risk in applying the solution of the old case to the current case, a new variable is introduced called the risk variable. This will reduce the influence of the weights on the overall evaluation of similarity and will also consider the not always positive effect of the solution of the selected case on our problem. This will make the retrieval stage better and more realistic.*
*Keywords: CBR, attribute risk, similarity measure, fuzzy inference system*

## 1 INTRODUCTION

Case-based reasoning (CBR) provides a methodology for reasoning and learning and there has been much published research on the subject (Jacek, 2001; Kolodmer, 1993; Roth-Berghofer, 2004). This methodology consists in using previous experiences (called cases) and in adapting previous solutions to solve new problems (i.e. it recalls old problems in order to obtain information about current problems).

In a CBR system, previous cases are stored in a case base and are characterized by a set of predefined attributes. When a new problem is encountered, the CBR system works through the following steps:

Retrieve similar cases to the new problem, reuse a solution suggested by a similar case, Revise the solution to fit the new problem, Retain the new solution once it has been confirmed.

Of these steps, case retrieval is extremely important for the success of a CBR system.

Although case-based reasoning is useful for various types of problems and domains (G. O. Yeh, 2001; Schmidt, 2001; Sadek, 2001; Long, 1995), it is not always the most suitable methodology to use. Below, we shall summarise some of the advantages and disadvantages of CBR.

Advantages:

1 Since CBR reduces the knowledge acquisition task, the human expert is then released from the task of giving training data.

2 The CBR system learns new cases when the CBR succeeds. These cases can be added to the case base and used to help solve future problems.

Disadvantages:

1 The CBR system requires a large amount of memory.
2 There is no successful technique for assessing the overall similarity between an old case and the target case within a case base.
3 The CBR system still finds it difficult to deal with uncertainty.

In this paper, we shall focus on this second point. It is of critical importance for there to be a good technique for retrieving cases in memory. If an appropriate case is not recovered from the case base, the rest of the process does not provide us with any useful results or information. Retrieval is a major research area in CBR, and research into the retrieval techniques used in CBR systems can be found in Watson (1999). All these techniques work in a similar way and involve developing a similarity metric. Examples of research into this subject can be found in Liao (1998) and Chen (1995).
The most commonly investigated retrieval techniques are those that employ the nearest-neighbor (NN) matching function. There are two steps to this technique: in the first step, the similarity of each individual attribute is calculated between the target problem and the old cases; and in the second, the overall similarity is calculated as the weighted sum of similarities between attributes. We shall illustrate this concept with an example. Let us consider the following situation. Two cases BALL and CAR are stored in a case base and we know what their categories are. A new case, CONSTRUCTION, is entered and we want to know what its category is.
The following table illustrates the cases, the value of the attributes, and the categories.

| TOY | BALL | CAR | CONSTRUCTION |
|---|---|---|---|
| Size | Medium | Very small | Very small |
| Principal Material | Plastic | Metal | Plastic |
| Small Pieces | No | Yes | Yes |
| Purpose | Fun | Fun | Fun |
| Players | 1 or more | 1 | 1 or more |
| Age | 2 | 3 | 2 |
| Place | Outside | Inside | Inside |
| SOLUTION | The category is OK | The category is DANGEROUS | ¿? |

*Table 1.*

We calculate the similarity between the cases BALL and CAR in relation to the current case CONSTRUCTION, and we shall see what category it is in. This is done in two steps. In the first step, we calculate the local similarity between attributes using the following local similarity function:

$$sim(x_i^{OLD}, x_i^{NEW}) = \begin{cases} 1 & si\ x_i^{OLD} = x_i^{NEW} \\ 0 & si\ x_i^{OLD} = x_i^{NEW} \end{cases} \qquad (1)$$

where $x_i^{OLD}$ is the $i^{th}$ attribute of the old case and $x_i^{NEW}$ is the $i^{th}$ attribute of the current case. In the second step, we calculate the overall similarity using the following arithmetic average equation:

$$Sim(C^{OLD}, C^{NEW}) = \frac{\sum_{i=1}^{n} sim(x_i^{OLD}, x_i^{NEW})}{n} \qquad (2)$$

where $C^{OLD}$ is the old case, $C^{NEW}$ is the target case, and $n$ is the number of attributes in each case. In order to avoid confusion, the overall similarity is referred to as $Sim$, and local similarity is referred to as $sim$.

Applying these two equations, we obtain the following result: $Sim(BALL, CONSTRUCTION) = 4/7$ and $Sim(CAR, CONSTRUCTION) = 4/7$.

It can be seen that there would be no difference between them since the measure does not consider many factors which are also important when cases are compared. The next step is to make the measure more accurate, and this is done by considering the importance of the attributes. In our problem, since some attributes are more important than others, it is therefore necessary to introduce the importance of the attributes as a new variable which we shall call the *weight variable*. This variable measures the importance of the $i^{th}$ attribute, and we shall write it as $\omega_i$.

| ATTRIBUTES | $\omega_i$ |
|---|---|
| Size | 0.08 |
| Principal Material | 0.24 |
| Small Pieces | 0.25 |
| Purpose | 0.21 |
| Players | 0.057 |
| Age | 0.09 |
| Place | 0.073 |

*Table 2.*

Using the weights associated to each attribute, we shall modify the overall similarity measure

$$Sim(C^{OLD}, C^{NEW}) = \sum_{i=1}^{n} \omega_i \cdot sim(x_i^{OLD}, x_i^{NEW}) \text{ , with } \sum_{i=1}^{n} \omega_i = 1 \qquad (3).$$

We calculate the similarity between the old case and the new case with the new equation and we obtain the following result:

$$Sim(BALL, CONSTRUCTION) = 0{,}597 \text{ y } Sim(CAR, CONSTRUCTION) = 0{,}613.$$

The value of the weights of the attributes is influenced by human subjectivity. Without changing the order of importance of the weights, small variations in the interpretation of the weights can cause the measure to choose one case or another. A simple example illustrates how small variations of the weights affect case retrieval. The table below shows the modified weights.

| ATTRIBUTES | $\omega_i$ | $\omega_i^{News}$ |
|---|---|---|
| Size | 0.08 | 0.077 |
| Principal Material | 0.24 | 0.242 |
| Small Pieces | 0.25 | 0.25 |
| Purpose | 0.21 | 0.203 |
| Players | 0.057 | 0.062 |
| Age | 0.09 | 0.096 |
| Place | 0.073 | 0.07 |

*Table 3*

Using the modified weights, we obtain these results:

$$Sim(BALL, CONSTRUCTION) = 0.603 \text{ and } Sim(CAR, CONSTRUCTION) = 0.6.$$

This result shows:

1. How the overall similarity measure is very sensitive to weight changes and how relevant the weights are when we retrieve a case, since as we can see with a small change in the weight, the measure recovers BALL as the most similar case to CONSTRUCTION. This cannot be right since the category of CONSTRUCTION is DANGEROUS (as it has small pieces) and not OK.
2. The weights are greatly influenced by human subjectivity.
3. The weights only consider the importance of the attributes.

4. The weights lose the local information obtained from measuring the risk in applying the solution of the old case to the current case. With this information, it will be easier to make a decision.

These are the reasons why we cannot put all the responsibility of the decision onto the weight.

In order to reduce the subjectivity of the weight assessments, we shall assign linguistic tags instead of accurate numbers and we shall consider a new variable that shall modify the measure indirectly rather than directly. This new variable shall quantify the risk of applying the old case solution to the target case, and shall contribute to the decision-making process. All the responsibility for the decision is no longer on the weight. With the weights and this new variable, we shall build a fuzzy inference system with which we shall calculate the similarity between the cases.

In this paper, we shall introduce a fuzzy inference system comprising 18 rules. The outputs of these rules is the measure presented in Section 2. In Section 3, we shall provide an example. Finally in Section 4, we shall outline our conclusions and future lines of work.

## 2 A NEW SIMILARITY MEASURE BASED ON A FUZZY INFERENCE SYSTEM

In order to improve the deficiencies of the weights, in addition to considering the importance of the weights, we shall also consider the difficulty of the solution of the retrieved case by applying it to the current case. We introduce a new variable called the *risk variable*. This variable shall quantify the danger taken when we apply the solution of the case in memory to the new case, and shall enable us to use the valuable local information that we obtain when we know whether the old case solution is useful for the target problem. This information will help us select the retrieval case more accurately.
As we saw in the example in Section 1, when the weight values were modified and we took BALL as the most similar case to CONSTRUCTION, the importance between attributes was taken into account, but we did not consider the risk we had by giving the solution OK to a dangerous toy: CONSTRUCTION had small pieces and was not safe for a child. In order to avoid this happening, we proposed the risk variable and for the $i^{th}$ attribute, this variable will be:

$$R_i \equiv \textit{The risk of applying the solution of the } C^{OLD}, \textit{when the } i^{th} \textit{ attribute of the } C^{NEW} \textit{ takes the value } A_i.$$

In our problem: if $C^{OLD}$ is BALL, $C^{NEW}$ is CONSTRUCTION and the solution of BALL is OK, the *risk* of applying OK-solution for the third attribute of CONSTRUCTION when the attribute takes the value *yes-small pieces* will be *High.* The new variable is introduced through a fuzzy inference system, and in order to build this system, we shall modify the formula of the overall similarity making it dependent on the risk ($R_i$) of the weight ($\omega_i$) and of the local similarity ($sim(x_i^{NEW}, x_i^{OLD})$).

$$Sim(C^{OLD}, C^{NEW}) = \sum_{i=1}^{n} f(R_i, \omega_i, sim(x_i^{OLD}, x_i^{NEW})) \qquad (4)$$

The function $f(\cdot)$ is implicitly obtained through a fuzzy inference system. We have chosen this system because it is able to work with imprecise or approximate information, and it is calculated as a weighted average of the outputs from all the rules. For each attribute, the local similarity is adjusted as an average of all the fired rules. The fuzzy inference system provides the following advantages in overall similarity assessment:

1 The local information obtained from the risk variable enables us to make the decision.
2 The decision is shared between the risk and the weights, and therefore its non-positive effect towards retrieval is alleviated.
3 Retrieval is better and more real.

The fuzzy inference system used in Equation 4 can be viewed as a variation of the TSK model, (Takagi, 1985). This inference system for the $i^{th}$ attribute contains nineteen rules as shown below.

Rule 1. If $R_i$ is *high* and $\omega_i$ is *high* and $sim(x_i^{OLD}, x_i^{NEW})$ is *high*,
then $v_i = \omega_i \cdot sim(x_i^{OLD}, x_i^{NEW}) - 0{,}1 \cdot (\omega_i \cdot sim(x_i^{OLD}, x_i^{NEW}))$.

Rule 2. If $R_i$ is *high* and $\omega_i$ is *high* and $sim(x_i^{OLD}, x_i^{NEW})$ is *medium*,
then $v_i = \omega_i \cdot sim(x_i^{OLD}, x_i^{NEW}) - 0{,}2 \cdot (\omega_i \cdot sim(x_i^{OLD}, x_i^{NEW}))$.

Rule 3. If $R_i$ is *high* and $\omega_i$ is *high* and $sim(x_i^{OLD}, x_i^{NEW})$ is *low*,
then $v_i = \omega_i \cdot sim(x_i^{OLD}, x_i^{NEW}) - 0{,}3 \cdot (\omega_i \cdot sim(x_i^{OLD}, x_i^{NEW}))$.

Rule 4. If $R_i$ is *high* and $\omega_i$ is *medium* and $sim(x_i^{OLD}, x_i^{NEW})$ is *high*,
then $v_i = \omega \cdot_i sim(x_i^{OLD}, x_i^{NEW}) - 0{,}2 \cdot (\omega_i \cdot sim(x_i^{OLD}, x_i^{NEW}))$.

Rule 5. If $R_i$ is *high* and $\omega_i$ is *medium* and $sim(x_i^{OLD}, x_i^{NEW})$ is *medium*,
then $v_i = \omega_i \cdot sim(x_i^{OLD}, x_i^{NEW}) - 0{,}3 \cdot (\omega_i \cdot sim(x_i^{OLD}, x_i^{NEW}))$.

Rule 6. If $R_i$ is *high* and $\omega_i$ is *medium* and $sim(x_i^{OLD}, x_i^{NEW})$ is *low*,
then $v_i = \omega_i \cdot sim(x_i^{OLD}, x_i^{NEW}) - 0{,}4 \cdot (\omega_i \cdot sim(x_i^{OLD}, x_i^{NEW}))$.

Rule 7. If $R_i$ is *high* and $\omega_i$ is *low* and $sim(x_i^{OLD}, x_i^{NEW})$ is *high*,
then $v_i = \omega_i \cdot sim(x_i^{OLD}, x_i^{NEW}) - 0{,}3 \cdot (\omega_i \cdot sim(x_i^{OLD}, x_i^{NEW}))$.

Rule 8. If $R_i$ is *high* and $\omega_i$ is *low* and $sim(x_i^{OLD}, x_i^{NEW})$ is *medium*,
then $v_i = \omega_i \cdot sim(x_i^{OLD}, x_i^{NEW}) - 0{,}4 \cdot (\omega_i \cdot sim(x_i^{OLD}, x_i^{NEW}))$.

Rule 9. If $R_i$ is *high* and $\omega_i$ is *low* and $sim(x_i^{OLD}, x_i^{NEW})$ is *low*,
then $v_i = \omega_i \cdot sim(x_i^{OLD}, x_i^{NEW}) - 0{,}5 \cdot (\omega_i \cdot sim(x_i^{OLD}, x_i^{NEW}))$.

Rule 10. If $R_i$ is *medium,* then $v_i = \omega_i \cdot sim(x_i^{OLD}, x_i^{NEW})$.

Rule 11. If $R_i$ is *low* and $\omega_i$ is *high* and $sim(x_i^{OLD}, x_i^{NEW})$ is *high*,
then $v_i = \omega_i \cdot sim(x_i^{OLD}, x_i^{NEW}) + 0{,}4 \cdot (\omega_i \cdot sim(x_i^{OLD}, x_i^{NEW}))$.

Rule 12. If $R_i$ is *low* and $\omega_i$ is *high* and $sim(x_i^{OLD}, x_i^{NEW})$ is *medium*,
then $v_i = \omega_i \cdot sim(x_i^{OLD}, x_i^{NEW}) + 0{,}3 \cdot (\omega_i \cdot sim(x_i^{OLD}, x_i^{NEW}))$.

Rule 13. If $R_i$ is *low* and $\omega_i$ is *high* and $sim(x_i^{OLD}, x_i^{NEW})$ is *low*,
then $v_i = \omega_i \cdot sim(x_i^{OLD}, x_i^{NEW}) + 0{,}2 \cdot (\omega_i \cdot sim(x_i^{OLD}, x_i^{NEW}))$.

Rule 14. If $R_i$ is *low* and $\omega_i$ is *medium* and $sim(x_i^{OLD}, x_i^{NEW})$ is *high*,
then $v_i = \omega_i \cdot sim(x_i^{OLD}, x_i^{NEW}) + 0{,}2 \cdot (\omega_i \cdot sim(x_i^{OLD}, x_i^{NEW}))$.

Rule 15. If $R_i$ is *low* and $\omega_i$ is *medium* and $sim(x_i^{OLD}, x_i^{NEW})$ is *medium*,
then $v_i = \omega_i \cdot sim(x_i^{OLD}, x_i^{NEW}) + 0{,}2 \cdot (\omega_i \cdot sim(x_i^{OLD}, x_i^{NEW}))$.

Rule 16. If $R_i$ is *low* and $\omega_i$ is *medium* and $sim(x_i^{OLD}, x_i^{NEW})$ is *low*,
then $v_i = \omega_i \cdot sim(x_i^{OLD}, x_i^{NEW}) + 0{,}1 \cdot (\omega_i \cdot sim(x_i^{OLD}, x_i^{NEW}))$.

Rule 17. If $R_i$ is *low* and $\omega_i$ is *low* and $sim(x_i^{OLD}, x_i^{NEW})$ is *high*,
then $v_i = \omega_i \cdot sim(x_i^{OLD}, x_i^{NEW}) + 0{,}3 \cdot (\omega_i \cdot sim(x_i^{OLD}, x_i^{NEW}))$.

Rule 18. If $R_i$ is *low* and $\omega_i$ is *low* and $sim(x_i^{OLD}, x_i^{NEW})$ is *medium*,

then $v_i = \omega_i \cdot sim(x_i^{OLD}, x_i^{NEW}) + 0{,}1 \cdot (\omega_i \cdot sim(x_i^{OLD}, x_i^{NEW}))$.

Rule 19. If $R_i$ is *low* and $\omega_i$ is low and $sim(x_i^{OLD}, x_i^{NEW})$ is *low*,

then $v_i = \omega_i \cdot sim(x_i^{OLD}, x_i^{NEW}) + 0{,}1 \cdot (\omega_i \cdot sim(x_i^{OLD}, x_i^{NEW}))$.

In these rules, the linguistic terms (*low*, *medium*, *high*) are defined on $R_i$, $\omega_i$, and $sim(x_i^{OLD}, x_i^{NEW})$, respectively; and $v_i$ $i = 1, \ldots, 19$ is the similarity measure adjusted by the system under the condition prescribed in the premise of the $j^{th}$ rule. The reason for the design of this fuzzy inference system is that the rules increase or decrease the influence of the local similarity on the overall similarity, depending on the risk, the weight, and the local similarity values.
Since there is high risk, high weight and high local similarity, there is a one-unit decrease in Rule 1; therefore, we cannot loose the good influence of them on the measure. There is a five-unit decrease in Rule 9 since an attribute with a high risk, and low weight and similarity is a bad influence on the overall similarity. There is a four-unit increase in Rule 11 because it is a very important and similar attribute with a low risk; it is very important in this context for the measure to be stronger. Rule 10 is neutral on the local similarity as we believe that if the risk is medium, it is not important enough to decide whether it should affect the measure. This simplifies our work since we need only consider critical cases.

## 3 A PRACTICAL EXAMPLE

In this section, an application of the proposed approach to the example given in Section 1 is illustrated. The first step for the application of the new measure is to assign linguistic terms to the weights and local similarities. The membership function is calculated for the weights and is shown below. The membership function is not calculated for the local similarity since we are dealing with qualitative rather than quantitative attributes. If the attributes were quantitative, then we could calculate the membership function in the same way as we did with the weights. In this example, the only values that take local similarity are 0 and 1 so we assign the following linguistic terms:

If $sim(x_i^{NEW}, x_i^{OLD}) = 1 \Rightarrow$ *High value.*

If $sim(x_i^{NEW}, x_i^{OLD}) = 0 \Rightarrow$ *Low value.*

The graph below shows the membership functions of the linguistic terms of the similarity measure.

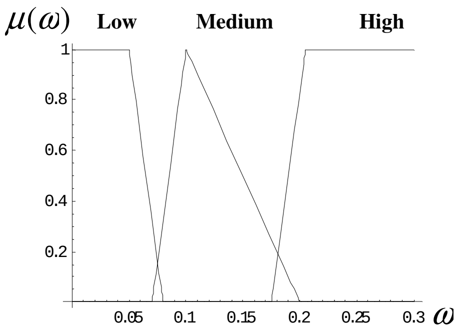


*Figure 1. Membership functions of the linguistic terms of the similarity measure*

The second step is to ask the human expert to assign the risk values. He/she must exclusively allocate risk to the critical cases (i.e. those cases with a high or a low risk). We are not interested in any other cases as they will not provide us with any useful information. The expert's task is also simplified by asking him/her for linguistic terms rather than accuracy numbers. The advantage of this is that it saves both time and effort.

| ATTRIBUTES | $R_i^{Ball,Const}$ | $R_i^{Car,Const}$ | $\omega_i$ | $\omega_i^{New}$ | *sim(Ball,Const)* | *sim(Car,Const)* |
|---|---|---|---|---|---|---|
| Size | High | Low | 0.08 | 0.077 | Low | High |
| P. Material | Medium | Low | 0.24 | 0.242 | High | Low |
| Small Pieces | High | Low | 0.25 | 0.25 | Low | High |
| Purpose | Low | Low | 0.21 | 0.203 | High | High |
| Players | Low | Low | 0.057 | 0.062 | High | Low |
| Age | Medium | Low | 0.09 | 0.096 | High | Low |
| Place | Low | Low | 0.073 | 0.07 | Low | High |

*Table* 4.

We shall now show how we apply the risk definition to some of the attributes.
$C^{OLD} = Ball$ , $C^{NEW} = Construction$ and $C^{OLD-SOLUTION} = OK$ .
First attribute = Size,
If Size = very small for construction and $C^{OLD-SOLUTION} = OK$ , the risk of applying solution OK for this attribute is High.
Second attribute = Principal Material,
If P. Material = plastic for construction and $C^{OLD-SOLUTION} = OK$ , the risk of applying solution OK for this attribute is Medium.

In the third step, we calculate the similarity. The following table shows the rules that the system fires.

| | | Attributes | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Ball | | | | | | | Car | | | | | | |
| | | A1 | A2 | A3 | A4 | A5 | A6 | A7 | A1 | A2 | A3 | A4 | A5 | A6 | A7 |
| Rules | Weights ( $\omega_i$ ) | 6 | 10 | 3 | 11 | 17 | 10 | 19<br>16 | 14 | 13 | 11 | 11 | 19 | 16 | 17<br>14 |
| | Weights ( $\omega_i^{New}$ ) | 6<br>9 | 10 | 3 | 11 | 17 | 10 | 19 | 14<br>17 | 13 | 11 | 11 | 19 | 16 | 17 |

*Table 5.*

In order to avoid confusion, $\omega_i$ similarity is denoted by $Sim$, and $\omega_i^{New}$ similarity by $Sim^{New}$. Overall similarity between the new case and Ball, and between the new case and Car is calculated by:

$$Sim(Ball, Construction) = 0 + 0.24 + 0 + (0.21 + 0.4 \cdot 0.21) + (0.057 + 0.3 \cdot 0.057) + 0.09 + 0 =$$
$$= 0.24 + 0.294 + 0.0741 + 0.09 = 0.6981.$$

$$Sim(Car, Construction) = (0.08 + 0.2 \cdot 0.08) + 0 + (0.25 + 0.4 \cdot 0.25) + (0.21 + 0.4 \cdot 0.21) + 0 +$$
$$+ 0 + (\frac{1}{2} \cdot [(0.073 + 0.3 \cdot 0.073) + (0.073 + 0.2 \cdot 0.073)]) = 0.096 +$$
$$+ 0.35 + 0.294 + \frac{1}{2} \cdot (0.0949 + 0.0876) = 0.83125$$

$$Sim^{New}(Ball, Construction) = 0 + 0.242 + 0 + (0.203 + 0.4 \cdot 0.203) + (0.063 + 0.3 \cdot 0.063) + 0.096 + 0 =$$
$$= 0.242 + 0.2842 + 0.0819 + 0.096 = 0.7041$$

$$Sim^{New}(Car, Construction) = \frac{1}{2} \cdot [(0.077 + 0.2 \cdot 0.077) + (0.077 + 0.3 \cdot 0.077)] + 0 + (0.25 + 0.4 \cdot 0.25) +$$
$$+ (0.203 + 0.4 \cdot 0.203) + 0 + 0 + (0.07 + 0.3 \cdot 0.07) =$$
$$= \frac{1}{2} \cdot (0.0924 + 0.1001) + 0.35 + 0.2842 + 0.091 = 0.82145$$

The above results show that Car is more relevant to the new case, Construction. If the attribute weights are changed and the overall similarity is recalculated, Car is still relevant to the new case. This proves the importance of the risk with regard to retrieval.

## 4 CONCLUSIONS AND FUTURE WORK

This article examines the deficiencies of similarity measures that only consider the weight factor and the non-positive influence that this may have on case recovery. In order to avoid such deficiencies, a new risk variable is proposed. This variable uses the local information obtained when the solution of the old case is applied to the value of the attribute leading to a better recovery. This also avoids the subjective influence of the weight on the measure. The result demonstrates that the new variable directly improves case retrieval. In order to enhance this method, we shall continue working on this line of research by studying a method to assign the risk.

## References

Jacek Jarmulak, Eugene J. H. Kerckhoffs and Peter-Paul van't Veen, 2001. Case-based reasoning for interpretation of data from non-destructive testing, *Engineering Applications of Artificial Intelligence* (14): 401-417.

J.L. Kolodmer, 1993. *Case-based Reasoning,* Morgan Kaufman, California.

Thomas R. Roth-Berghofer, 2004. Explanations and Case-Based Reasoning: Foundational Issues, *Lecture Notes in Computer Science.*

Anthony G. O. Yeh and Xun Shi, 2001. Case-based reasoning (CBR) in development control, *International Journal of Applied Earth Observation and Geoinformation* (3): 238-251.

Rainer Schmidt and Lothar Gierl, 2001. Case-based reasoning for antibiotics therapy advice: an investigation of retrieval algorithms and prototypes, *Artificial Intelligent in Medicine* (23): 171-186.

Adel W. Sadek, Brian L. Smith and Michael J. Demetsky, 2001. A prototype case-based reasoning system for real-time freeway traffic routing, *Transportation Research Part C: Emerging Technologies*: (9): 353-380.

Long, D.; King, R.L.; Lack, 1995. R., Control of power systems using case-based reasoning, S*ystem Theory, Proceedings of the Twenty-Seventh Southeastern Symposium*: 73-74.

I. Watson, 1999. Case-based reasoning is a methodology not a technology, *Knowledge-Based System* (12): 303-308.

T. Warren Liao, Zhiming Zhang and Clauder R. Mount, 1998. Similarity measures for retrieval in Case-based reasoning systems, *Applied Artificial Intelligence* (12): 267-288.

Shy-Ming Chen, Ming-Shiow Yeh, Pei-Yung Hsiao, 1995. A comparison of similarity measures of fuzzy values, *Fuzzy Sets and System* (72): 79-89.

T. Takagi and M. Sugeno, 1985. Fuzzy identification of systems and its application to modeling and control, *IEEE Transactions on Systems, Man, and Cybernetics* (15): 116-132.