Using Fuzzy Integral to Model Case-Base Competence

Simon C. K. Shiu, Yan Li, Xi Z. Wang
Department of Computing
Hong Kong Polytechnic University
Hung Hom, Kowloon, Hong Kong
{csckshiu, csyli, csxzwang}@comp.polyu.edu.hk

Abstract

Modeling case-base competence is a crucial issue in the field of CBR. In fact it has led to a number of important developments in CBR recently, the most well known being in case deletion and case-base visualization and authoring support. This is highlighted by Smyth and Keane's competence model (called the S-K model in this paper), which was shown to be efficient. In this paper, we argue that the model is not always a good predictor of competence, especially when problem distributions are non-uniform. Consequently, a new competence model based on fuzzy (non-linear) integral is proposed to address such a problem. The main idea is to repartition competence groups to ensure that the distribution of each group is nearly uniform, and more importantly to use the fuzzy integral with respect to a fuzzy measure (non-additive set function) to compute the overall competence. The interaction among new competence groups is considered to be reflected in the non-additive set function. The advantage of the newly proposed model is its high accuracy for predicting competence especially in the situation of non-uniform distributed case-base.

1 Introduction

In the field of artificial intelligence, case-based reasoning (CBR) is an effective and efficient problem-solving technique by reusing the solutions to similar problems stored as cases in a case-base [1]. With the growing use of the CBR system, there is a surge of research on case-base maintenance (CBM), whose aim is to achieve optimal performance. Case-base competence (coverage), the range of target problems that can be successfully solved, is a critical factor contributing to the performance of a CBR system. The idea that one can accurately model the competence of a case-base is a powerful one. In fact, it has led to a number of important applications in the area of CBM, especially in competence-based CBM [2], which focuses on increasing the overall competence of the case-base.

The most frequently used competence model is proposed by Smyth and Keane [3]. In their model, statistical properties such as the size and density of cases, as well as the problem-solving properties of the given CBR system such as retrieval and adaptation characteristics have been taken into account. However, the more important idea in the model is to regard a competence group instead of a single case as the fundamental unit of competence. In fact, the definition of competence group implies that different groups have no interaction (overlap) with each other. The model was empirically shown to be competent and efficient.

In this paper, we present a detailed analysis of the S-K competence model and show that it does not always work well on real problems since it assumes a uniform distribution of case-bases. We argue that in the situation of non-uniform distributed case-bases, the S-K model gives a bad prediction of case-base competence. In response, we have developed a different competence model that is based on fuzzy integrals (non-linear integrals) with respect to a fuzzy measure (or a non-addition set function). The main idea is to repartition competence groups to ensure that the distribution of each group is nearly uniform, and more importantly to use fuzzy integrals with respect to a fuzzy measure (a non-additive set function) to compute the overall competence. The model developed in this paper consists of four steps. First, we identify competence groups in a given case-base according to the S-K model. Second, we formulate the concept of non-uniform distribution (or uneven distribution), which is related to case-base coverage. Third, we seek weak links in every given competence group, and repartition the competence groups. Fourth, we use fuzzy integrals with respect to a fuzzy measure (or a non-additive set function) to compute the overall coverage of a given case-base. The interaction among the new competence groups is considered to be reflected in the non-additive set function. Compared with the S-K model, the advantage of our competence model is that we can deal with more general case-bases, especially non-uniform distributed case-bases.

2 Related work

For the artificial intelligence community especially CBR systems, the competence or coverage of a given system is a fundamental evaluation criterion. Recently, the issue of competence has received much attention from the perspective of the so-called case-base maintenance problem, that is, the issue of how best to manage the organization and contents of a case-base in order to optimize future reasoning performance.

The importance of case competence has been brought sharply into focus since many maintenance policies are directly linked to heuristics that attempt to measure case competence to guide maintenance procedures [4-7]. However, these competence heuristics have provided only coarse-grained estimates of competence. For example, Smyth and Keane employed a category-based competence model that classifies cases as belonging to one of only four possible competence categories; Zhu and Yang described case coverage based on a rather rough concept of case neighborhood.

The most recently explicit algorithmic model of competence for case-based reasoning systems was suggested by Smyth and Keane [3]. Several innovative solutions to problems based on their model have been developed, such as the construction of compact competent case-bases [2], the case retrieval problem [8], and the provision of case authoring support and case-base visualization [3].

3 Analyzing the competence model

In this section, a detailed analysis of Smyth and Keane's case-base competence model is presented. It shows that the original model does not always work well in real world problems.

3.1 A Review of Case-base Competence

The competence of a case-base system (the range of problems it can solve) depends critically on the cases in the case-base, but the relationship between cases and overall competence is very complex. To address this problem, Smyth and Keane suggest a competence model based on the concept of competence group, which implies that different groups have no interaction (overlap) with each other. In their model, both statistics and problem-solving properties are considered. Two key involved fundamental concepts are coverage and reachability. Coverage of a case refers to the set of problems that the case can solve. Reachability is the set of cases that can be used to provide solutions for a problem. And to characterize the competence of a case-base in a tractable fashion, a reasonable assumption is made, which considers the case-base a representative sample of the problem space. With this technique, a given case-base can be partitioned to several different competence groups. The competence of a competence group G (group coverage) depends on the number and density of cases in G (see Equation 1).

$$GroupCoverage(G) = 1 + [|G| \cdot (1 - GroupDensity(G))] \quad (1)$$

where group density is defined as the average case density of the group (see Equation 2).

$$Groupdensity(G) = \frac{\sum_{c \in G} CaseDensity(c, G)}{|G|} \quad (2)$$

The local density of a case c within a group of cases G is considered to be the average similarity of c to any other cases in the group G.

For a given case-base, with competence groups $G = \{G_1, G_2, ..., G_n\}$, the total coverage is defined by Equation 3.

$$Coverage(G) = \sum_{G_i \in G} GroupCoverage(G_i) \quad (3)$$

It obvious from Equation 2 that the S-K model assumes a uniform distribution. In other words, the more uniform the cases distributed in the group, the more suitable the competence model for the given case-base. Here, uniform distribution means that the case density for each case in a competence group is the same. However, for many case-bases in the real world, this condition is too strict to satisfy. Hence, the model is not always accurate, especially in the situation of non-uniform distributed case-bases.

3.2 Non-uniform Distribution

We can show that the S-K model is not a good predictor when the case-base is non-uniform distributed. Here, we only need to give an example to illustrate this point. Suppose that, in some domain, we have the graph structure as shown in Figure 1.

To make the point clear, in Figure 1, suppose that the densities of group G1 and group G2 are both 0.8, and are both assumed to be uniform-distributed (i.e. the case density of each case in either G1 and G2 respectively, is 0.8) while the density of the whole group $G = G1 \cup G2 \cup \{c^*\}$ is 0.2, and the coverage of c^* is 3 cases, but the overlap coverage of c^* and $G1 \cup G2$ is 2, and c^* is a pivotal case. Intuitively, we can straightforwardly obtain the coverage of the whole competence group G as follows:

$$GroupCoverage1(G) = GroupCoverage(G1) +$$
$$GroupCoverage(G2) + [coverage(c^*)$$
$$- coverage(c^*) \cap coverage(G1 \cup G2)]$$
$$= 1 + [|G|(1 - GroupDensity(G))] + 1 + [|G2|(1 - GroupDensity(G))] + 1$$
$$= 1 + 5(1 - 0.8) + 1 + 7(1 - 0.8) + 1$$
$$= 5.4$$

However, according to the above competence model, the results are as follows:

$$GroupCoverage2(G) = 1 + [|G| \cdot (1 - GroupDensity(G))]$$
$$= 1 + 12(1 - 0.2)$$
$$= 10.6$$

We can see that the two results are very different from each other. Since G1 and G2 fit the competence model perfectly, the coverage of G1 and G2 can be computed using the S-K model, so the GroupCoverage1 should perform well according to our conditions given in advance. The problem is in the process of computing GroupCoverage2, which assumes that group G is also uniform-distributed, resulting in a computing error that cannot be ignored.

In cases such as Figure 1, the difference between GroupCoverage1 and GroupCoverage2 (which can be regarded as the computing error of the S-K model) is as follows:

$$\begin{aligned}
& GroupCoverage2 - GroupCoverage1 \\
&= \{1 + [|G| \cdot (1 - GroupDensity(G))]\} \\
&\quad - \{1 + [|G1| \cdot (1 - GroupDensity(G1))] + 1 \\
&\quad + [|G2| \cdot (1 - GroupDensity(G2))] + 1 \\
&= |G| \cdot [(GroupDensity(G1) - GroupDensity(G)] \\
&\quad - 1 - GroupDensity(G1) \quad \ldots\ldots(4) \\
&\geq |G| \cdot [(GroupDensity(G1) - GroupDensity(G)] - 2 \quad \ldots(5)
\end{aligned}$$

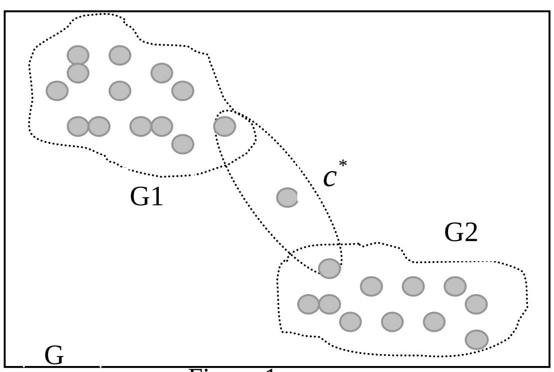

Figure 1

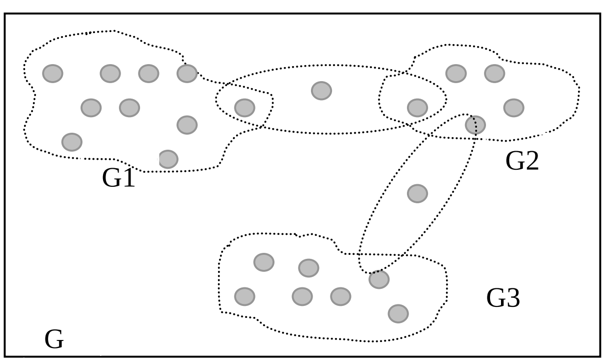

Figure 2

According to the above results, if we let the number of cases (which satisfy given conditions in competence group G1 or G2) tend to ∞, $[GroupDensity(G1) - GroupDensity(G)]$ increases at the same time, and as a result, the computing error tends to ∞. It is the group density difference between G1/G2 and G which causes this competence error.

4 Fuzzy integral competence model

In the example, we see that for a non-uniform distribution, the S-K model is not a good predictor of case-base competence. It's reasonable to compute the competence of G1, G2 and c^* respectively. The competence of the original competence group G is considered to be their summation. In fact, we can easily see that case c^* has an important role in given competence group G. It is the case that mainly affects the overall distribution of a competence group, and further affects the precision of computing overall competence using the S-K model. Thus, the idea is very natural for recursively detect such cases as c^* (which can intuitively be called weak-links) and repartitioning the competence group G to several smaller groups such as G1 and G2, whose distributions are regarded to be uniform. Their competence can then be computed by using the S-K model. The competence of the weak-links can be considered to be their individual coverage respectively, which reflects the relationship among the several new groups.

However, the conditions given in the example are too restricted to be satisfied for a common competence group. More generally, new groups $G1$ and $G2$ are not necessarily strictly uniform-distributed, and the weak link case c^* is not necessarily a pivotal case. To deal with this situation, without influencing the results in Equation 4, $GroupDensity(G1)$ can be replaced by the average group density of group G1 and G2, which can be denoted by $\overline{GroupDensity(G_i)}$ $i \in \{1,2\}$,so $[GroupDensity(G1) - GroupDensity(G)]$ is equal to $[\overline{GroupDensity(G_i)} - GroupDensity(G)]$ denoted by $\Delta GroupDensity$.

We then introduce a concept called quasi-uniform distribution to describe the distribution closed to uniform distribution, which relaxes the condition of strict uniform distribution of each group. Consider another assumption that c^* is a pivotal case in the example, this is not necessarily true in many cases. To address this problem, just consider the individual competence of c^* as its relative coverage, which is defined as follows (Equation 6):

$$RelativeCoverage = \sum_{c' \in CoverageSet(c)} \frac{1}{|ReachabilitySet(c')|} \quad (6)$$

Hence, according to Equation 4 and 5,

$$\begin{aligned} & Competence-error(c^*) \\ & = |G|\Delta GroupDensity - \overline{GroupDensity(G_i)} - \\ & \mathrm{Re}\,lativeCoverage(c^*) \\ & \geq |G|\Delta GroupDensity - (\mathrm{Re}\,lativeCoverage(c^*) + 1) \end{aligned} \quad (7)$$

According to inequality 7, since $\mathrm{Re}\,lativeCoverage(c^*)$ is limited, so we can see that it is $\Delta GroupDensity$ which leads to the competence error. In other words, it is the assumption of uniform distribution of the competence group that makes the performance of the S-K model turn bad in some situations.

4.1 Detecting weak-links

To improve the predicting performance of the S-K model, what we should first to do is to identify the weak-links in each competence group in the S-K competence model. The next task is to compute the overall coverage of the given competence group. To complete this task, an explicit definition of weak-link should be given in advance.

Since our aim is to predict case-base competence (or coverage) more accurately, in this paper, the definition of weak-link, as well as several relative concepts, is more directly related to the competence of the group in question.

Def. Let $G = \{c_1, c_2, ...c_n\}$ be a given competence group in a case base $C, c^* \in G$, is called a weak-link if

$$\begin{aligned} & CompetenceError(c^*) = |G|\Delta GroupDensity - \\ & \overline{GroupDensity(G_i)} - \mathrm{Re}\,lativeCoverage(c^*) \\ & \geq \alpha \end{aligned}$$

where α is a parameter which is different according to different requests. If $\exists c^* \in G$, c^* is a weak-link, then the competence group G is called a non-uniform distributed competence group. Otherwise, if for $\forall c \in G, CompetenceError(M) \leq \alpha$, then G is called a quasi-uniform distributed competence group.

It's obvious that $\sum G_i = G - \{c^*\}$. With the definition, we propose the recursive method to detect weak-links in a given competence group G, which is described by the following algorithm:

Weak-link Detection Algorithm:

1.$W-SET \leftarrow \{\}, G-SET \leftarrow \{\}, i = |G|$;

2. If ($i \neq 0$)

{Consider each given competence group G in the S-K competence model, compute $CompetenceError(c)$, $\forall c \in G;\ i = i-1$}

3. If there is no weak-link, add G to $G-SET$, end;

4.If there is a weak-link c^*, identify the competence groups $G_1, ..., G_n, (n \geq 1)$ in $G - \{c^*\}$ using the S-K competence model, and add c^* to the set of weak-links *W-SET*.

5. For $(1 \leq i \leq n)$

{$G \leftarrow G_i$; repeat step1 to step4}.

Thus, we can obtain the set of weak-links $W-SET$ in a given competence group G and the set of new competence groups $G-SET$. Obviously, a given competence group G is repartitioned by identifying weak-links in it. The groups in $G-SET$ are called new competence groups.

4.2 Computing overall coverage of the given competence group using fuzzy integral

After detecting weak-links in a competence group G, several new competence groups $G_1, ..., G_n (n \geq 1)$ are produced. According to the definition of a weak-link, each new produced group is sure to be quasi-uniform distributed. The next task is to compute overall coverage of G. In the example described in Figure 1, we just add the competence of $G_i (1 \leq i \leq n)$ and the relative coverage of c^*, but this method has no representativeness. There are more complicated situations, one of which is illustrated in Figure 2. It's difficult to clearly identify the contribution of each weak-link. In fact, in Figure 2, c^* has much more influence on coverage of G than c^{**} has, which reflects different relationship among new competence groups. To describe the complex relationship, a powerful tool called fuzzy integral (or non-linear integral) with respect to a fuzzy measure (a non-additive set function) is applied.

4.2.1 Non-additive set function

Let X be a nonempty set and $\Phi(\mathrm{X})$ be the power set of X. We use the symbol μ to denote a non-negative set function defined on $\Phi(\mathrm{X})$ with the properties $\mu(\Theta) = 0$. If $\mu(\mathrm{X}) = 1, \mu$ is said to be regular. It is a generalization of classic measure [9]. When X is finite, μ is usually called a fuzzy measure if it satisfies monotonicity, i.e.,

$$\mathrm{A} \subseteq \mathrm{B} \Rightarrow \mu(\mathrm{A}) \leq \mu(\mathrm{B})\, for\, \mathrm{A}, \mathrm{B} \in \Phi(\mathrm{X})$$

For a non-negative set function μ, there are some associated concepts. μ is said to be additive if $\mu(\mathrm{A} \cup \mathrm{B}) = \mu(\mathrm{A}) + \mu(\mathrm{B})$ for $\mathrm{A}, \mathrm{B} \in \Phi(\mathrm{X})$;to be sub-additive if $\mu(\mathrm{A} \cup \mathrm{B}) \leq \mu(\mathrm{A}) + \mu(\mathrm{B})$ for $\mathrm{A}, \mathrm{B} \in \Phi(\mathrm{X})$; to be super-additive if $\mu(\mathrm{A} \cup \mathrm{B}) \geq \mu(\mathrm{A}) + \mu(\mathrm{B})$ for $\mathrm{A}, \mathrm{B} \in \Phi(\mathrm{X})$.

Let $\mathrm{X} = \{G_1, ..., G_n\}$ be the space of the new competence groups, A and B are two subsets of the power set of X. Here, A, B respectively can be a single new group G_i or the union of several groups. If we consider $\mu(\mathrm{A})$ as the importance of subset A, then the additivity of the set function means that the joint importance of the several groups is just the sum of their respective importance, which implies that there is no interaction among competence groups. However, this is not true in the considered problem. In fact, most measures of importance are non-additive.

Sub-additivity and super-additivity are two special types of non-additivity. Super-additivity means that the joint importance of two sets is greater than or equal to the sum of their respective importance, which indicates the two sets enhance each other. In contrast, sub-additivity means that the joint importance of two sets is less than or equal to the sum of their respective importance, which indicates that the two sets resist each other.

In our problem, consider $\mathrm{X} = \{G_1, ..., G_n\}$ to be the factor space. There are weak-links among the competence groups which link them to one group G. Here, weak-links such as c^* and c^{**} are sure to enhance the overall coverage of G. Hence, the importance measure μ defined on the power set $\Phi(\mathrm{X})$ is a super-additive measure. So here we have

$$\mu(\mathrm{A} \cup \mathrm{B}) \geq \mu(\mathrm{A}) + \mu(\mathrm{B}) \text{ for } \mathrm{A}, \mathrm{B} \in \Phi(\mathrm{X})$$

For example, in Figure 2, c^* enhances the importance of $G_1 \cup G_2$, c^{**} enhances the contribution of $G_1 \cup G_3$, and there is no case to enhance or reduce the contribution of $G_2 \cup G_3$, so we have

$$\mu(G_1 \cup G_2) \geq \mu(G_1) + \mu(G_2)$$
$$\mu(G_1 \cup G_3) \geq \mu(G_1) + \mu(G_3)$$
$$\mu(G_2 \cup G_3) = \mu(G_2) + \mu(G_3)$$

Using fuzzy integral to compute the overall coverage of the original competence group G, the importance measure μ should be determined in advance. However, for a factor space including n factors, there are ($2^n - 1$) parameters to decide [10]. In the situation of Figure 2, 7 values of the importance measure should be determined, say, $\mu(G_1)$, $\mu(G_2)$, $\mu(G_3)$, and $\mu(G_1 \cup G_2)$, $\mu(G_1 \cup G_3)$, $\mu(G_2 \cup G_3)$, $\mu(G_1 \cup G_2 \cup G_3)$ should be given.

To reduce the load, we apply a kind of fuzzy measure called $\lambda -$ *fuzzy* measure, which is in the following form:

$$\mu(\mathrm{A} \cup \mathrm{B}) = \mu(\mathrm{A}) + \mu(\mathrm{B}) + \lambda \cdot \mu(\mathrm{A}) \cdot \mu(\mathrm{B}) \quad \lambda \in (-1, \infty)$$

if $\lambda \leq 0$, μ is a sub-additive measure; if $\lambda \geq 0$, μ is a super-additive measure; if and only if $\lambda = 0$, μ is additive. So here we have $\lambda \geq 0$. Applying $\lambda -$ *fuzzy* measure to determine the importance measure μ, what we need to do is just to determine the n importance on each single factor and λ.

4.2.2 Determining $\lambda -$ *fuzzy* measure μ

In this paper, we consider that the importance of each competence group is equal to 1, i.e. $\mu(G_i) = 1, (1 \leq i \leq n)$. This assumption is reasonable because each group makes unique contribution to the overall coverage, that is, the status of each group is considered to be equal.

Next task is to determine the parameter λ, which is critical to determine μ. It's obvious that the properties of the weak-links between the two groups are important factors for determining λ. In our model, coverage of a group refers to the area of the target problem space covered by the group. In this sense, the value of λ is closely related to the coverage of the weak-links and the density of their coverage sets. Consider arbitrary two new groups G_i and G_j, the $W-SET$ between them is $C^* = \{c_1^*, ..., c_h^*\}$.

Define $Coverage(C^*)$ and $Density(C^*)$ as follows:

$$Coverage(C^*) = \sum_{i=1}^{h} RelativeCoverage(c_i^*)$$

$$Density(C^*) = \sum_{i=1}^{h} GroupDensity(Cov(c_i^*)) / h$$

where $Cov(c_i^*)$ is the coverage set of one of the weak-links between G_i and G_j.

The coverage contribution of $G_i \cup G_j$ must be directly proportional to $Coverage(C^*)$ and inversely proportional to $Density(C^*)$. With these preparations, the parameter λ is given by the formula in Equation 8.

$$\lambda = Coverage(C^*) \cdot (1 - Density(C^*)) \qquad (8)$$

Then the $\lambda -$ *fuzzy* measure μ is determined.

4.2.3 Using Choquet integral to compute overall group competence

Due to the non-additivity of the set function μ, some new types of integrals (known as non-linear integrals) have to be used, which can be considered to be the generalization of the weighted mean [9]. The advantage of using fuzzy integrals is that the interactions of all factors in a factor space can be taken into account. Fuzzy integrals have found a few applications in CBR systems [11-12]. X. Z. Wang and D. S. Yeung used fuzzy integrals to compute the overall similarity between problems and each stored case in the case-base for each feature, interactions among the features are considered. In a Cash Flow Forecasting system, Rosina Weber Lee et al. made use of fuzzy integrals in the same way to choose the best match for a problem in the retrieved cases.

A common type of nonlinear integral with respect to non-negative monotone set functions is the Choquet integral.

Let f be a non-negative real-valued measurable function defined on X, and μ be a non-negative monotone set function introduced in the above section. The Choquet integral of f on X with respect to μ, (c) $\int f d\mu$, is defined by the formula

$$(c)\int f d\mu = \int_0^\infty \mu(F_\alpha) d\alpha,$$

where $F_\alpha = \{x | f(x) \ge \alpha$ for any $\alpha \in [0, \infty)$. When X is finite, the Choquet integral can also be defined in the same way with respect to a non-negative set function that is not necessarily monotone.

In our model, $X = \{G_1, ..., G_n\}$ is finite, $f_i = GroupCoverage(G_i)$, importance measure μ satisfies:

$\mu(G_i) = 1 (1 \le i \le n)$;

$\mu(A \cup B) = \mu(A) + \mu(B) + \lambda \cdot \mu(A) \cdot \mu(B)$ ($\lambda \ge 0$)

where λ is determined by Equation 8.

The process of calculating the value of Choquet integral is as follows:

(1) rearranging $\{f_1, f_2, ..., f_n\}$ into a non-decreasing order such that

$$f_1^* \le f_2^* \le \cdots \le f_n^*$$

where $(f_1^* \le f_2^* \le \cdots \le f_n^*)$ is a permutation of $(f_1, f_2, ..., f_n)$;

(2) computing

$$(c)\int f d\mu = \sum_{j=1}^{n} [f_j^* - f_{j-1}^*] \cdot \mu(\{G_j^*, G_{j+1}^*, \cdots, G_n^*\})$$

where $f(x_0^*) = 0$.

The value of Choquet integral is considered as the coverage of the considered competence group. Each competence group in the S-K model is considered in the same way, the sum of all group coverage is the overall coverage of the given case-base.

5 Simulation

In previous sections, a new model using fuzzy integral for modeling the competence of case-bases has been presented. In this section, empirical evidence is needed to support this model. In short, we demonstrate that the model proposed in this paper closely match the actual competence. At the same time the S-K model is shown not to be a good predictor when the case-base is not uniform distributed.

In this section, we use a small case-base, which contains 120 cases. Each case is chosen randomly such that the case-base satisfies non-uniform distribution. Every case is a two-dimension vector. For experimental reasons, 50 randomly chosen cases in the case-base are used as unseen target problems, a further 70 cases are used to form the experimental case-bases.

In our experiment, the success criterion used is a similarity threshold: if the system does not retrieve any cases within this threshold, a failure is announced. True competence is regarded as the number of successfully solved problems.

The experiment is repeated 100 different times, and the average results are computed which are shown in Table 1. The results positively support our model.

We use “Error_number”, “Error_percentage” as evaluate indexes, which represent the relative error of coverage computed by using the S-K model and the fuzzy integral model(new model) respectively. Here, Error_percent = Error_number /True_competence

Index	True	S-K model	New model
Density	-	0.4	0.6
Competence	34.5	49.6	38.9
Error_number	0	15.1	4.4
Error_percent	0	43.8%	12.8%

Table 1.

Clearly the results are very positive. The error_percentage of our fuzzy integral model is rather lower than using the S-K model. When the number of cases increases, our model can strikingly reduce the competence error compared to the S-K model.

We should also point out that in our experiment, the case-base considered has a non-uniform distribution, but in the situation of uniform distributed case-bases, the fuzzy integral competence model can still be used. Because if there is no weak-link, the competence computed by the fuzzy integral model is equal to the results of the S-K model, which has been proved to be effective. Thus, the uniform distributed case-base is a special case for our model.

6 Conclusions and Discussion

In this paper, a novel competence model of case-bases is proposed. A type of fuzzy integral (called choquet integral) with respect to the $\lambda - fuzzy$ measure is used as a tool. We consider the non-uniform distributed case-bases, and prove that the competence model proposed by Barry Smith and Keane is not a good predictor in such a situation. The results of experiments conducted are very positive for our model. It shows that the new model proposed in this paper has extended the scope of modeling case-base competence.

7 Acknowledgement

This project is supported by the Hong Kong Polytechnic University research grants G-T144 and H-ZJ90.

References

[1] J.Kolodner,Case-Based Reasoning, Morgan Kaufmann: San Mateo, CA, 1993.

[2] B.Smyth and E.McKenna, Building Compact Competent Case-Bases, Case-Based Reasoning Research and Development. Lecture Notes in Artificial Intelligence (Althoff, K-D, Bergmann, R. and Branting, L, K. eds.), Springer Verlag, 1999, pp. 343-357.

[3] B.Smyth and E.McKenna, Modeling the Competence of Case-bases, Advances in Case-Based Reasoning. Lecture Notes in Artificial Intelligence, pp.208-220, 1998

[4] B.Smyth and E.McKenna, Remember to Forget: A Competence Preserving Case Deletion Policy for CBR Systems, Proceedings of the 14th International Joint Conference on Artificial Intelligence, pp.377-382, 1995

[5] B.Smyth, Case-based Maintenance, Proceedings of the 11th International Conference on Industrial **&** Engineering Applications of AI Expert Systems, Springer Verlag.

[6] J. Zhu and Q. Yang, Remembering to Add: Competence-preserving Case-Addition Policies for Case-Base Maintenance. Proceedings of the 16th International Joint Conference on Artificial Intelligence, 1999, pp. 234-239.

[7] D.Leake and D.Wilson, Remembering Why to Remember: Performance-Guided Case-base Maintenance, Advances in Case-based Reasoning. Lecture Notes in Artificial Intelligence, pp.161-172

[8] B. Smyth and E. McKenna, Footprint-Based Retrieval, Case-Based Reasoning Research and Development. Lecture Notes in Artificial Intelligence (Althoff, K-D, Bergmann, R. and Branting, L. K. eds.), Springer Verlag, 1999, pp. 329-342.

[9] Z. Wang and G.J.Klir, Fuzzy Measure Theory, Plenum, New York, 1992

[10]Z. Wang, K.S. Leung and J. Wang, A genetic algorithm for determining non-additive set functions in information fusion, Fuzzy Sets and Systems. Vol.102, pp. 463-469, 1999.

[11]X. Z. Wang and D.S.Yeung, Using Fuzzy Integral to Modeling Case-based Reasoning with Feature Interaction, 2000 IEEE International Conference on Systems, Man, and Cybernetics, Vol. 5, 2000, pp.3660 –3665

[12]W. L. Rosina, M. B. Ricardo, and K. K. Suresh, proceedings of the first international conference of Case-based Reasoning, ICCBR-1995, (eds.) Manuela Veloso and Agnar Aamodt, Springer Verlag, pp. 510-519.

[13]K.Racine, and Q. Yang, Maintaining Unstructured Case. Proceedings of the 2nd International Conference on Case Based Reasoning, pp. 553-564 RI, USA.