# Using a Two-Layered Case-Based Reasoning for Generating Scientific Techniques

MAHSA KIANI* [1], KAMBIZ BADIE** , MARYAM TAYEFEH MAHMOUDI **

*Faculty of engineering, Computer Group
Islamic Azad University, Science and Research Campus
Hesarak, Poonak, Tehran
IRAN

**Info-Society Department
Iran Telecom Research Center (ITRC)
North of Kargar Ave., Tehran
IRAN

*Abstract: -* In this paper, a two-layered case-based reasoning has been discussed for adjusting the degrees of importance for the attributes of problem context, which are to be used for retrieving similar cases. More over, weighted average technique is used to perform compositional adaptation on the values of some generic attributes, which can be used to represent mathematical functions with different formats. The proposed approach can be workable enough for integrating the previously-experienced techniques to handle a new problem context.

*Key-Words:* Technique generation, Case-based reasoning, Case retrieval, Compositional adaptation, Secondary case-based reasoning, Weighted average

## 1 Introduction

An aspect of design or planning is generating new models or techniques, which can have a lot of applications in research and development type problems. In this respect, a previous research work has been done, showing how new scientific models/techniques can be generated based on composing the models or techniques already experienced in the past [1]. The approach was demonstrated to be quite workable for the situations, where an integration of previously-experienced techniques has something to contribute to a new situation. In this paper, as an extension of the approach represented in [1], weighted average technique is used to perform compositional adaptation on the values of some generic attributes, which can be used to represent mathematical functions with different formats. More over, 2ndary case-based reasoning is used to adjust the degrees of importance for the attributes that are used in the process of retrieving the similar cases from the case-library.

## 2 Some Previous Works in the Domains of Design, Planning and Generating New Models/Techniques

### 2.1 Some Previous Works in the Domain of CBR-Based Design & Planning

There are some systems which apply substituational adaptation for design purposes. For example, STEPPC has been developed as a generic tool to design process tracing and reuse, based on a substitutional adaptation for the parameters [2]. With respect to transformational adaptation, MRL system is a domain-independent case based planner, which reuses plans based on deductive planning approach [3]. A CBR-based system has also been proposed, which is able to make design patterns under constraints for engineering systems in general,

[1] Any correspondence to:
{Mahsa_Kiani@yahoo.com}

and electronic systems in particular, based on a sort of transformational adaptation that takes into account the status of incompatibility of an old solution toward the new problem context [4]. Regarding derivational adaptation, PARIS is a domain independent case-based planning system in which abstract case is retrieved and refined by a generative planner [5]. Finally, regarding compositional adaptation, FAMING [6] has been proposed for designing part shapes in mechanisms. Meanwhiles, COMPOSER, which is used in engineering design, applies constraint satisfaction algorithm for adaptation [7]. There is also a Tutoring Library System that generates a new set of chapters to provide a new book [8].

### 2.2 A Previous Work in the Domain of CBR-Based Generation of New Models / Techniques

Because of complexity in problems situation, it is not necessarily feasible to develop a sound model or technique for a new situation based on predefined analytical principles. In a previous work [1], authors proposed an approach, based on composing the techniques suggested earlier for sufficiently similar situations. According to this approach, having retrieved sufficiently similar cases based on some similarity measures, the tree-like structures related to the corresponding solutions, are composed hierarchically such that the values belonging to the similar generic actions beneath each node can be regarded for composition. If these values are some parameters, which are not furtherly expandable in the tree, composition is done using the technique of minimizing the summation of the distances between the final solution and a case solution, otherwise the same process will be repeated for the values of the generic actions beneath them. Also, with respect to optimization of sequences of action which is an essential stage within the compositional adaptation process, a global distance function is utilized. The entire process will be repeated until all the nodes in the solution trees were considered for composition. The approach has been tested for generating new techniques in the domain of Image Processing, and results of adaptation have been quite satisfactory. The techniques located at the nodes of tree solution, however may indicate different mathematical expressions, thus leading to a situation where compositional adaptation can not be performed simply at parameter level. Due to this fact, human expert was introduced as an interactive means for giving the final idea on total composition of these mathematical expressions. In this paper, as an extension of the previous work, we make use of weighted average technique as a means for composition of the values discussed above. In the meantime, since the quality of the cases to be retrieved depends on the degrees of importance for each attribute, it may be necessary to determine the related values on the ground of past experiences. For this purpose, the application of a previously-developed technique called secondary case-based reasoning [9] is discussed.

## 3 The Proposed Complementary Process of Technique Generation Based on A Two-Layered Case-Based Reasoning

### 3.1 Basics

The quality of cases which are to be retrieved is related to the degree of importance for each attribute and these values are not the same. Let say, values of the degrees of importance are related to the problem context, and these parameters should therefore be re-valued based upon the detected context. In this respect, a secondary CBR is used. Each case in the secondary CBR, contains a situation which is the problem context, and a solution that is suitable values for degrees of importance for the attributes. Regarding this, the primary CBR cycle tries to generate a suitable technique for the current problem context, and the secondary CBR cycle tries to determine the bias values for retrieval phase of the primary CBR. The adaptation process of secondary CBR is performed in compositional manner. Since the solutions in primary CBR are the techniques with tree-like structure, it is important to consider them in a way that they can get ready for adaptation. Regarding this, those nodes in the solution tree which follow a similar action, are composed in their technique value. These techniques may have different expressions and could not be composed directly. So techniques sholuld be expressed in such a way that can have similar attributes. In this respect, each technique could be expressed in terms of some structural specificities. Weights that are used in this technique, are normalized similarity between each case situation and the current problem context. In addition, some techniques are presented in terms of mathematical functions. Because these functions usually have different expressions, they can not be composed directly. To similaritize their mathematical structures, each function is expressed by means of its "Taylor expansion". On this basis, each mathematical function is expressed as a sum of variables with the exponent signs and its coefficients. Coefficients with similar exponent of the variables, are then composed taking into account the normalized similarity between the current problem context and the corresponding case situation as the weight. In case of no similirity between the functions domains or their structural specifications, the human expert, will be asked to give the final idea on the appropriate techniques.

### 3.2 Process of Technique Generation Based on Two-Layered Case-Based Reasoning and Composing Mathematical function with Different Formats

First, past experiences of researchers in generating techniques are represented in terms of appropriate cases including case situation and case solution. In addition, successful experiences in determining suitable values for degrees of importance for the attributes are stored in a secondary CBR.

Facing a new problem, appropriate values for the degrees of importance for each attribute in the current problem context, are first determined by a secondary CBR [9]. For this purpose, similar cases are retrieved from the secondary case base and their solutions are composed with respect to the normalized similarity of each case situation to current problem context. The values of parameters are considered as weights for computing the normalized distance between current problem context with situation of stored cases in primary CBR. The mathematical expression in this regard is as follows:

$$Case\_Distance_i = \frac{\sum_{j=1}^{m} (Dist(A_i^j, A^j).W^j)}{\sum_{i=1}^{m} W^j} \quad (1)$$

$Case\_Distance_i$ is the normalized distance between current problem context with situation of $Case_i$ in primary case base. $Dist(A_i^j, A^j)$ is the distance of values for the $attribute_j$ in situation of $Case_i$ with current problem context. $W^j$ is the value for the degree of importance for the $attribute_j$ in the current problem context, which is determined by a secondary CBR.

Here, selected cases are those with a normalized distance as well as a variance less than certain thresholds. Techniques which are the solutions of these selected cases, should be then adapted. The process of adaptation goes back to three different compositional adaptation, for adapting the values of coefficients of “Taylor expansion” related to terminal techniques, for adaptation of structural specificities value of techniques for similar actions, and finally for adapting sequences of actions justifying the related techniques.The two first stages of adaptations are performed by weighted average technique taking into account the normalized similarity between the corresponding case situation and the current problem context as weight. The last stage of adaptation is done by applying a global distance function [10].

## 4 An Example in the Domain of Image Processing

Suppose that we would like to propose a suitable technique for situation which is shown in Fig. 1(a). To do so, regarding a suitable threshold for assessing distance between current problem context and situation of each existing case in secondary case library, suitable similar cases will be selected. In this example four cases with distances of “0.143” , “0.1714” , “0.257” and “0.271” are selected and their solutions which are the values of degrees of importance for each attribute in the case situation, are composed in proportion of normalized similarity of each case situation with current problem context. The normalized similarity of these cases are “0.857”, “0.8286”, ”0.743” and “0.729” respectively. Finally the importance value of each attribute in the current probem situation, respectively are: “7.407”, ”8.235”, “9.222”, “7.996”, “3.618”, “5.488” and “5.999”. Then these values of parameters are considered as weights for distance between values of the attributes in each case situation and the current problem context. Finally, normalized distance value between current problem context and situation of existing cases in primary case library is going to be computed.

In this example threshold is considered as “0.3” and four cases of “No. 4”, “No. 11”, “No.10” and “No.1” respectively with normalized distances of “0.121”, “0.126”, “0.194”, and “0.286” are selected as similar cases. Then their normalized similarity respectively are “0.879”, “0.874”, “0.806” and “0.714”. These cases along with their solutions are shown in Fig. 1(b), 1(c), 1(d) and 1(e).

In order to have a suitable technique for the situation of current problem, techniques with similar actions are being composed and this process is repeated until the stage where a technique is represented in terms of some parameters values which can become simply subject to composition, or some predefined structures whose composition can be carried out by a human expert who is in interaction with the system. For instance, in Fig.1 four soution trees have actions of “feature extraction” and then “assessment”, their techniques are composed together then technique which is applied for “feature extraction” includes of “assessment” with composing “Normalized co-occurrence matrix in four different direction (joint probability)”, “Feature assessment (Reflectivity magnitude AND correlation coefficient)”, “Intensity value assessment” and “First order statistical analysis (energy) Daubechies wavelet” with normalized similarity of cases as weight. Techniques used for “classification” in second layer of case “No. 4”, “No. 11”, “No. 10” are now composed together and as three of them are neural network and have similar structural specificities, the values of their parameters composed in proportion of normalized similarity between each case situation and current problem context. Their structural specificities are consist of number of layers, average number of neurons, bias unit value, learning rate and momentum factor. For example, Momentum factor in case "No.4", "No.11", "No.10" are "0.6", "0.6" and "0.9" respectively. Composition of them with respect to normalized similarity of cases is "0.6945". The primary values of them and the composed results are shown in Table 1. In addition in three cases, ("No. 4", "No. 11", "No. 10"), "classification, classification and mapping" actions, are allowed to be composed. The techniques for "mapping" action are three different mathematical activation function that are represented in terms of Taylor expansion. Coefficients which have similar order of variables, should have been composed regarding normalized similarity between each case situation and current problem context as weight. For example, coefficient of "x" in taylor expansion for case "No.4", "No.11", "No.10" are "0.5", "0.75" and "0.25" respectively. Composition of them with respect to normalized similarity of cases is "0.5066".Primary functions and composed resulted functions are shown in Table 2.

| | |
|---|---|
| MLP structure 1 (Case NO.4) | Structure= (6, 7, 1.0, 0.05, 0.6) |
| MLP structure 2 (Case NO.11) | Structure= (4, 9, 0.03, 0.03, 0.6) |
| MLP structure 3 (Case NO.10) | Structure= (3, 10, 0.1, 0.01, 0.9) |
| MLP Compositional structure | Structure= (4,9,0.385,0.03057, 0.6945) |

Table 1 Composition of Structures **[**Structure=(Number of layers, Average number of neurons, Bias unit value, Learning rate, Momentum factor)**]**

| | |
|---|---|
| Activation function 1 (Case NO.4) | $f(x) = 0.5 + 0.5x - 0.166x^3$ |
| Activation function 2 (Case NO.11) | $f(x) = 0.5 + 0.75x - 0.562x^3$ |
| Activation function 3 (Case NO.10) | $f(x) = 0.5 + 0.25x - 0.0208x^3$ |
| Compositional . . . . function | $f(x) = 0.5 + 0.5066x - 0.2555x^3$ |

Table 2 Composition of Functions

Then adaptation for obtaining optimized sequences of actions must be done. By minimizing the global distance function, optimal architecture of solution tree for the current problem context is illustrated in Fig. 1(a). The ratios indicated in the figure, stand for the normalized similarities of the corresponding retrieved cases with respect to the current problem context.

The textual expression for the final solution of Fig. 1(a) is as follows:

… The appropriate proposition of technique for the current problem of classification, whose context is already described in Fig. 1(a), is respectively, to perform "preprocessing", "feature extraction" and next to that "preprocessing" based on its own techniques, and finally to perform "classification" with the "MLP" technique.

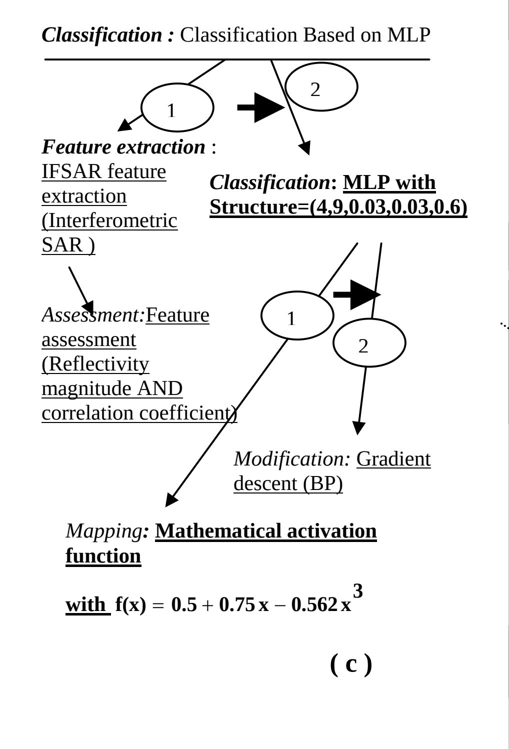
Phase of "preprocessing", is based on "enhancement" by "filtering", in this respect, composit "transformation" technique should be applied from "Histogram stretch transformation" with "Structure=(Number of bins in histogram=20)" and "Gaussian filter (transformation)". Next phase is "feature extraction" which is performed by composing the "Normalized co-occurrence matrix in four different direction (joint probability)", "Feature assessment (Reflectivity magnitude AND correlation coefficient)", "Intensity value assessment" and "First order statistical analysis (energy) Daubechies wavelet". The following phase is "preprocessing" which is done by "feature selection". Technique for doing "feature selection" is "SOM (self organizing maps)" which its structural specificities are "Structure=(Network size=64, Learning rate =0.9, Error threshold =0.05, Initial value for gain function in winning cluster=0.3, Initial value for gain function in neighborhood border=0.05)". "SOM (self organizing maps)" has two generic actions: "selection" and "modification" which is done by "Euclidean distance" and "Weight modification (Learning)", respectively. Finally "classification" in the second layer of solution tree is performed by "MLP" technique. Structural specificities for this technique is "Structure=(Number of layers=4, Average number of neurons=9, Bias unit value=0.385, Learning rate=0.03057, Momentum factor=0.6945)". To perform "MLP" technique, two actions are necessary: "mapping" with $\mathbf{f(x) = 0.5 + 0.5066\,x - 0.2555\,x^3}$ as "compositional mathematical activation function" and "modification" with "Gradient descent (BP)" which is the same in three cases.

It is seen that the values of parameters for case retrieval in the primary CBR are assigned by using the secondary CBR Then most suitable cases for composition will be retrieved and finally most appropriate technique will be generated for the current problem context. Therefore, through a compositional adaptation of the pre-stored cases, which are sufficiently similar to the current problem context, we will be able to generate a novel technique with appropriate values in its structural specificities, with based on the phases which are

**Goal:** Classification

**Entity**: Gray Scale image
**Status of existing knowledge**: Vector
**Status of environmental noise**: -------------
**Status of processing time:** --------------
**Status of flexibility in answer:** ------------
**Technique previously used:** Based on usual image Descriptors (edge information)
**Misperformance perceived:-** Low accuracy

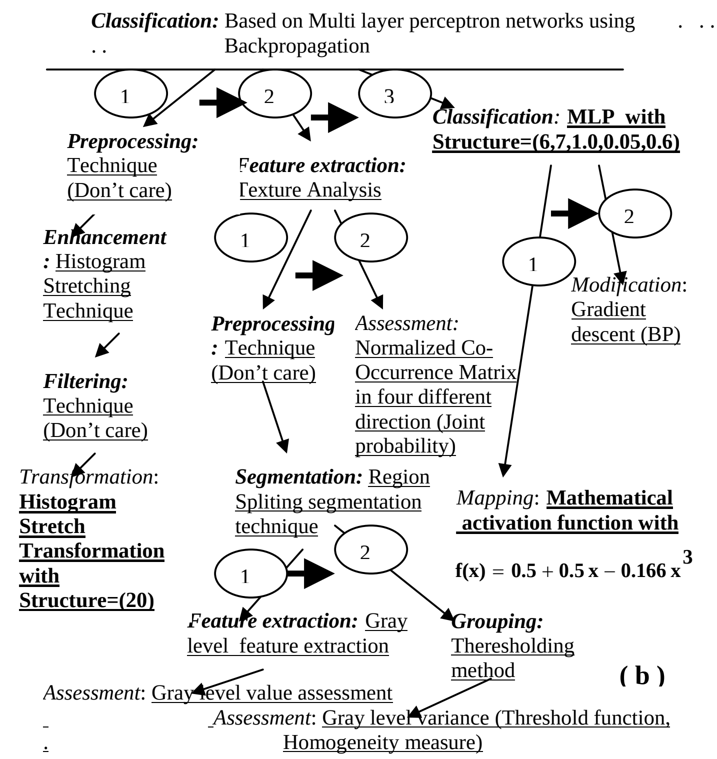
Case NO .11 ( Retrieved Case 2 ) *PROBLEM CONTEXT* — *SOLUTION*




**Goal:** Classification

**Entity**: Gray Scale image
**Status of existing knowledge**: Vector
**Status of environmental noise**: ------------
**Status of processing time:** --------------
**Status of flexibility in answer:** Medium flexibility
**Technique previously used:** Non-Linear bilateral subtraction technique with gray level thresholding
**Misperformance perceived:** Low accuracy

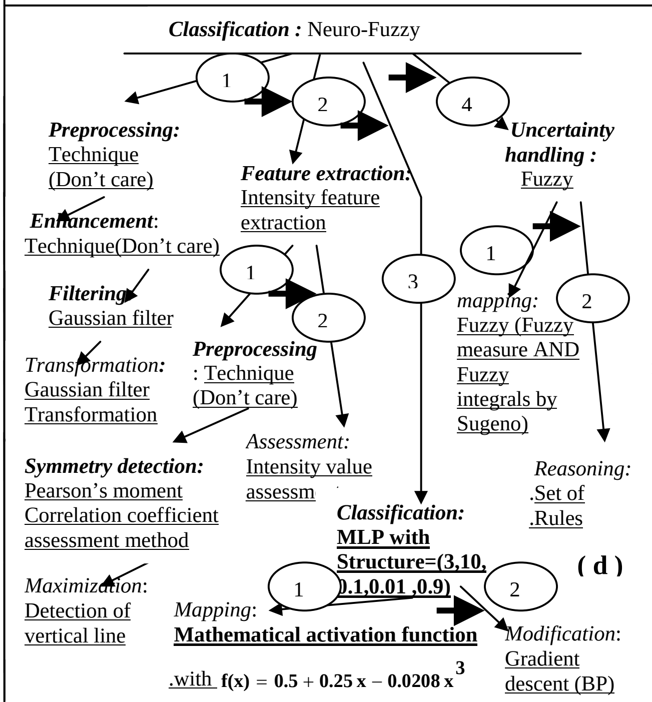
Case NO 4 ( Retrieved Case 1 ) *PROBLEM CONTEXT* — *SOLUTION*




**Goal:** Classification

**Entity**: Gray Scale image
**Status of existing knowledge**: Vector
**Status of environmental noise** ---------------
**Status of processing time:** ---------------
**Status of flexibility in answer:** Low flexibility
**Technique previously used:** Neural Network
**Misperformance perceived:** Low accuracy

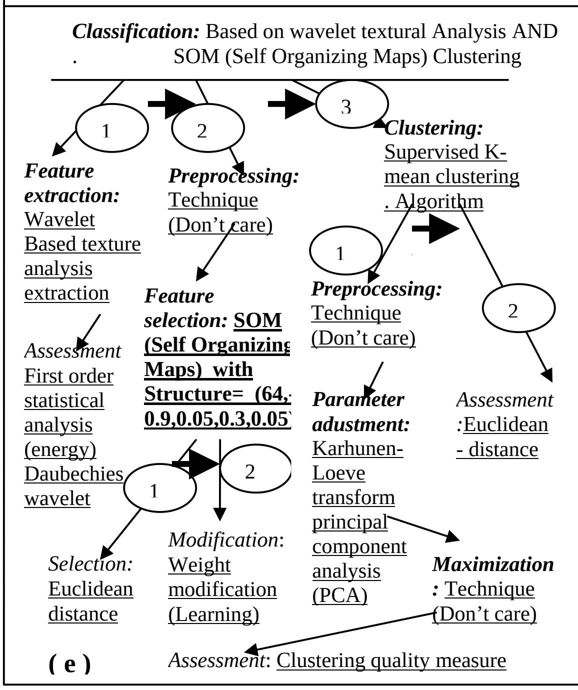
Case NO .10 ( Retrieved Case 3 ) *PROBLEM CONTEXT* — *SOLUTION*




**Goal:** Classification

**Entity**: Gray Scale image
**Status of existing knowledge**: Vector
**Status of environmental noise**: Optical noise
**Status of processing time**: ----------
**Status of flexibility in answer**: -----------
**Technique previously used**: Based on usual descriptors(gray-level)
**Misperformance perceived**: Low precision

Case NO .1 ( Retrieved Case 4 ) *PROBLEM CONTEXT* — *SOLUTION*



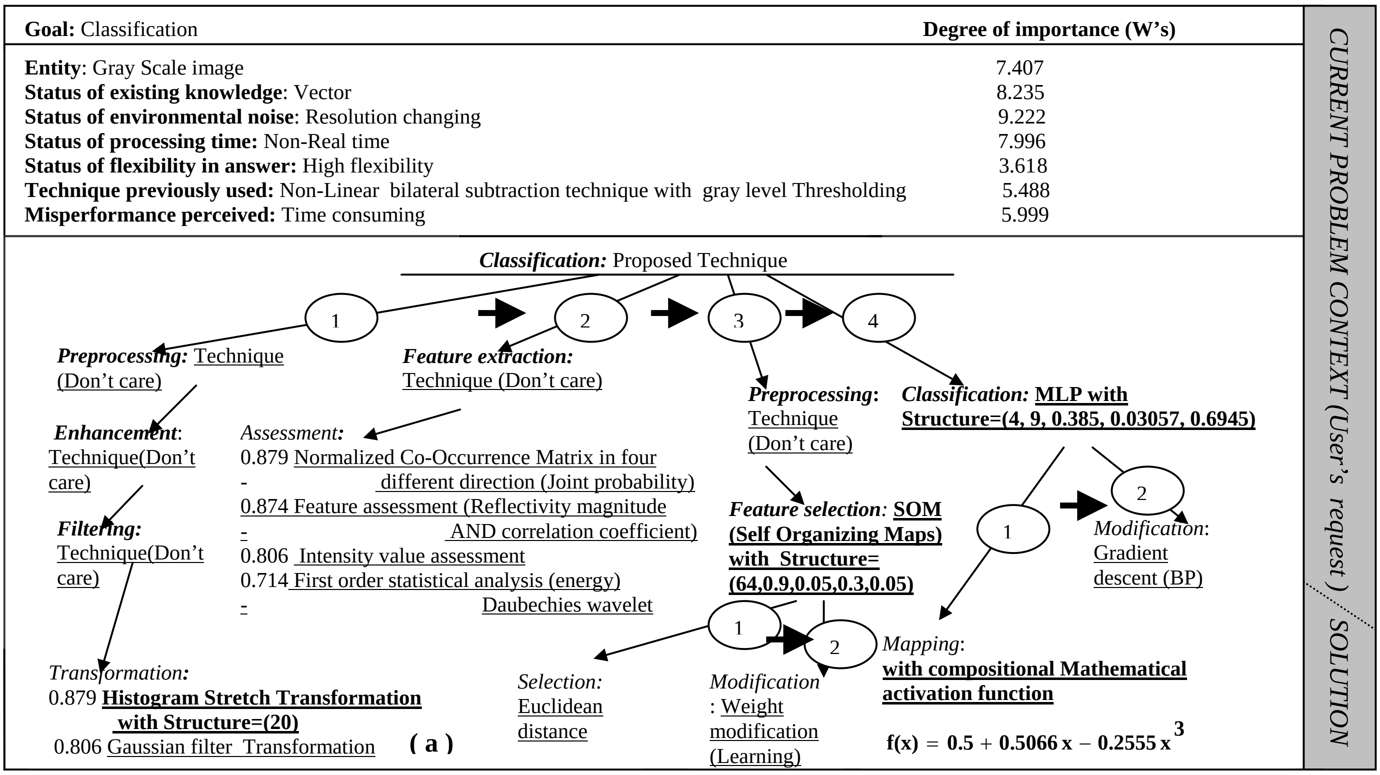



Fig.1 An example

described in a sequential manner.

## 5 Concluding Remarks

It was discussed that, weighted average technique can be used as a suitable means for composing the values of the attributes which are able of representing mathematical techniques/expressions with different formats. Meanwhiles, it was demonstrated that a secondary case-based reasoning system can help obtain efficient values for the degrees of importance regarding the attributes in problem context. It is anticipated that experience management via the approach discussed in this paper, can help much generation of the techniques, which are somehow similar to those by expert researcher / innovators.

*References:*


[1] K.Badie, M.T.Mahmoudi and M.Kiani, An approach to generating scientific techniques based on compositional adaptation , *Proc. of 9th Annual CSICC'04*, Vol. II, pp.55-63, Iran, 2004.

[2] N.Prat, Applying case-based reasoning in a design environment: issues and lessons learned, *6th Intl. Conf. on Artificial Intelligence in Design (MLinD'00)*, Massachusetts, June 2000.

[3] J.Koehler, Flexible plan reuse in a formal framework, *Current Trends in AI Planning, Proc. 2nd European Workshop on AI Planning (EWSP'93)*, pp.171-184, 1994.

[4] M.Hejazi and K.Badie, A new approach to solution adaptation and its application for design purposes, *6th European Conf. (ECCBR'2002)*, Vol. 2416, pp.549, 2002.

[5] R.Bergmann and W.Wilke, Building and refining abstract planning cases by change of representation language, *Journal of Artificial Intelligence Research*, No.3, pp.53-118, 1995.

[6] B.Faltings, FAMING: Supporting innovative design using adaptation: A description of the approach, implementation, illustrative example and evaluation, *Engineering Design Synthesis*, 2002.

[7] L.Purvis and P.Pu, COMPOSER: A case-based reasoning system for engineering design, *Robotica*, Vol.16, No.3, pp. 285-295, 1998.

[8] N.Arshadi and K.Badie, A compositional approach to solution adaptation in Case-Based Reasoning and its application to tutoring library, *Proc. of 8th German Workshop on Case Based Reasoning*, Lammerbuckel, 2000.

[9] K.Badie and N.Reyhani, Learning through estimating optimal formats for problem solving, *IEEE Intl. Conf. on Systems & Man and Cybernrtics*, Hammemet, Tunisia, 2002.

[10] N.Reyhani, K.Badie and M.Kharrat, A new approach to compositional adaptation based on optimizing the global distance function and application in an intelligent tutoring system, *IEEE Intl. Conf. on Information Integration & Reuse*, 2003.