Aesthetical Decision Making of Music Composer Using ANFIS

M. Reza Mashinchi, M. Hadi Mashinchi, Ali Selamat

Faculty of Computer Science& Information Systems, Universiti Teknologi Malaysia, Johor, Malaysia
r_mashinchi@yahoo.com, *h_mashinchi@yahoo.com*, *aselamat@fsksm.utm.my*

Abstract

One aesthetical face that human is strongly surrounded by it, is music, which we are going to clarify it here. For this purpose, fuzzy logic gives us the possibility of organizing our view on this matter, due to its uncertainty and flexibility. Fuzzy rules and fuzzy reasoning are the backbone of fuzzy inference systems, which are the most important modeling tools in fuzzy set theory. They are often employed to capture the imprecise modes of reasoning that play an essential role in human ability to make decisions in environment having uncertainty and imprecision. In this paper, we use fuzzy inference system as an appropriate reasoning tool. A class of adaptive networks that are functionally equivalent to fuzzy inference systems is referred as ANFIS. We will try to use ANFIS for implementation modeling of decision making for music composing, to obtain final composed notes.

1. Introduction

The problem of knowledge understanding and representation become central in computer sciences. The problem formulation and solving, understanding human requests in natural languages, machine learning and other urgent problems of computer simulation are faced to understanding the entire diversity of human knowledge. Throughout the centuries human is trying to know aesthetic and its rules, and so the philosophers have pondered over the questions related to it [5]. Actually human does trying everything to be on this aesthetic rules. Despite of the human awareness, wanted or unwanted, his or her doing is on this way and it is scaled by these rules [2]. Therefore, due to the above discussion, understanding of decision making via aesthetic cognition is strongly recommended.

Everything human is doing is to get a goal, and reaching to this goal has a satisfaction that can be a kind of aesthetic itself. However, aesthetic is consequently a reminder to a beauty, which is always coherence with joy. But in fact, it has caused many obscurities and problems in human life. In another meaning, it has made aesthetic foe itself [5], due to our wrong understanding and reasoning. To explain common sense reasoning, Craik [6] viewed the brain as a system for making models: "If the organism carries a small-scale model of external reality and of its own possible actions within its head, it is able to carry out various alternatives, conclude which is the best of them, react to future situations before they arise, utilize the knowledge of past events in dealing with the present and the future and in every way react in a fuller, safer, and more competent manner …". To simulate such a system, Minsky [9] proposed the notion of frames, which are prefabricated patterns, assembled to form mental models. But on the other hand, if we refer aesthetic to common sense, up to now no one has produced exact description of its meaning and maybe it will be never happen. Therefore, by considering it as a vague concept, for analyzing, we need to have an appropriate tool. Such one that is invented in 1965 by professor L.A. Zadeh started with speaking on beauty as ordinarily concept, he was debating with his friend the issue of which wife is more beautiful [10]. The meaning of fuzzy is unvocal, vague, giddy and inexact, which helps us to have understanding and dealing better with vagueness of aesthetic. Therefore, it leads us to understand more of the harmonies in aesthetics, which is depended on its each single member's harmony and also that is relative to each person individually [3,8]. In this paper we propose a model of an external reality in composing a music that will support the best reactions of the listeners on alternatives. Also we will identify some notions of assumed melodies which are prefabricated to form our final composition by using fuzzy inference system as a way of reasoning to reach our goal to compose a musical prediction.

2. Aesthetic cognition

Aesthetic is referred to anything related to beauty that its description is ambiguous [5]. Therefore due to vagueness of aesthetic, human have problem to rule it by specified scales. In aesthetic expression, we are always using linguistic variables to describe and valuate this beauty. By using the linguistic hedges like "very", "more", "less" or any other ones, the beauty could be described clearly. As a result, we have identified that a fuzzy logic which deals with vague concepts of these linguistic hedges is an appropriate tool to be applied here. As it is quoted in [10], explained by L. A. Zadeh, conventional techniques for system analysis are intrinsically unsuited for dealing with humanistic systems, whose behaviors is strongly related to human judgment and perception. This is a manifestation of what might be called the principle of incompatibility: "as the complexity of a system increases, our ability to make precise and yet significant statements about its behavior diminishes until a threshold is reached beyond which precision and significant become almost mutually exclusive characteristics"[10]. It was because of this believe that Zadeh proposed the concept of linguistic variables [10] as an alternative approach to modeling human thinking, an approach which has an approximate manner, serve to summarize information and express it in terms of fuzzy sets instead of crisp numbers. As we are trying to achieve human/animal realization [4] full of complexity, we can model aesthetic cognition explained in the next section.

2.1. Aesthetic modeling

To model the cognition of aesthetic we explain the AMP, whose three important components are RW, AW and GA as in Figure 1. In the first part RW, we have all ambiguous and unknown raw data. In the second AW, it is simple and trusted part and the third one is GA, which is guessed aesthetic. For the detail refer to [3].

In an attempt to apply AMP for music composer [2,8], we show it as in Figure 2.

In the first part, it is coming all sounds to hear for a composer that is difficult to rule it and so it will be changed to some simple and trustable data that would be resulted in the second part. To make prediction of the scene the third part of our model is needed to have this result. It is in this part that he will compos music noted of the scene.

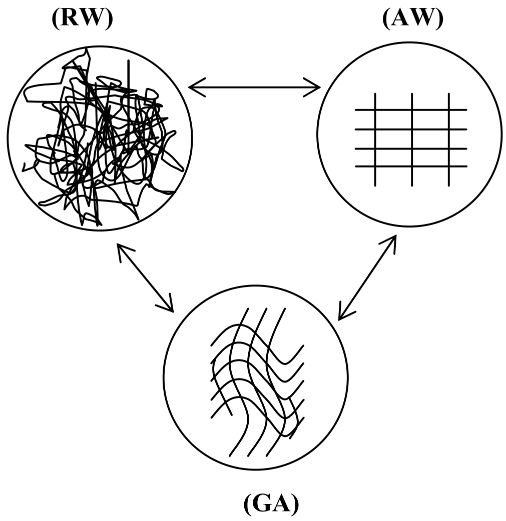

Figure 1. Aesthetic Modeling Process (AMP)

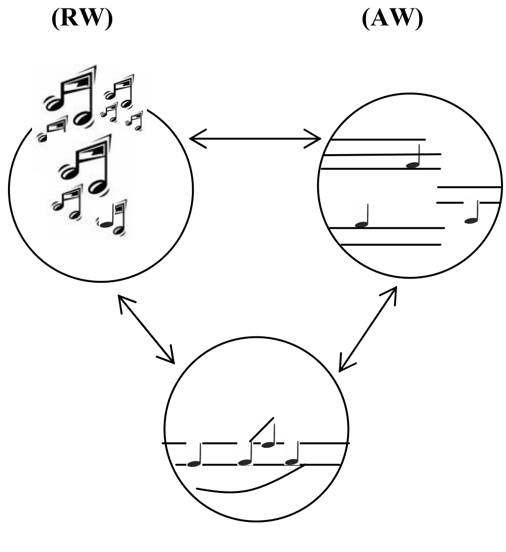

Figure 2. AMP for Music composing

Now, Fuzzy Inference Rule Base (FRIB) [3] answers this question that how the trusted and predicted data are obtained from ambiguous data on second and third part. There we have rules to make changes on raw data.

3. ANFIS implementation on music composer modeling

In this section, we review some materials [7] that we are going to use in the rest of the paper. A class of adaptive networks that are functionally equivalent to fuzzy inference systems is referred as ANFIS standing for adaptive network based fuzzy inference system [7], see figure 6. We can implement the model AMP for music composer by ANFIS. For decision making in AMP in Figure 2, we have to rule data in the second and the third parts. So, due to course of dimensionality as examples, we will choose only some few rules of all from FIRB. We attempt to implement by two input, a fist-order Sugeno fuzzy model with nine rules, where each input is assumed to have three associated membership functions (MFs). We assume two input to be able to use grid partitioning. Our linguistic variables are characterized by a quintuple $(x, T(x), G, M)$ in witch x is the name of the variable, $T(x)$ is the term set of x, that is the set of its linguistic values or linguistic terms; X is the universe of discourse; G is a syntactic rule which generates the terms in $T(x)$; and M is semantic rule which associate with each linguistic value A its meaning $M(A)$, where $M(A)$ denotes a fuzzy set in X.

Now by considering Example 1 of [3], $T(x_1)$ and $T(x_2)$ can be obtained as:

$T(sword)$: {very strong, not strong, strong, not very strong,…., medium strong…..}

$T(whine)$: {very high, not high, high, not very high,…, medium high,…., low,…}

which each term in $T(sword)$ and $T(whine)$ is characterized by fuzzy set of discourse $X = [0,100]$, where the above hedges are explained in the following. Let A be a linguistic value characterized by a fuzzy set with membership function $\mu_A(.)$. Then A^k is interpreted as a modified version of the original linguistic value whose membership function is:

$$\mu_{A^k}(x) = \int_X [\mu_A(x)]^k / x. \tag{3-1}$$

In particular, the operation of concentration is defined as

$$CON(A) = A^2 \tag{3-2}$$

while that of dilation is expressed by

$$DIL(A) = A^{0.5} \tag{3-3}$$

Conventionally, we take $CON(A)$ and $DIL(A)$ to be the result of applying the hedges very and more or less, respectively, to the linguistic term A.

We can interpret the negation operator NOT and connectives AND and OR as:

$$NOT(A) = \neg A, \tag{3-4}$$

where

$$\mu_{\neg A}(x) = \int_X [1 - \mu_A(x)] / x, \tag{3-5}$$

$$A\ AND\ B = A \cap B, \tag{3-6}$$

where

$$\mu_{A\cap B} = \int_X [\mu_A(X) \wedge \mu_B(x)] / x, \tag{3-7}$$

$$A\ OR\ B = A \cup B, \tag{3-8}$$

$$\mu_{A\cup B} = \int_X [\mu_A(X) \vee \mu_B(x)] / x, \tag{3-9}$$

respectively, where A and B are two linguistic values whose meaning are defined by $\mu_A(.)$ and $\mu_B(.)$, where x is the given input, with the interval [0,100] as the universe of discourse. Then we can construct MFs for the following composite linguistic terms as follow:

- More or less war $= DIL(war) = war^{0.5}$

$$= \int_X \sqrt{\frac{1}{1+(\frac{x-100}{30})^6}} \Bigg/ x \tag{3-9}$$

- Not war and not peace $= \neg war \cap \neg peace$

$$= \int_X \left[1 - \frac{1}{1+\left(\frac{x}{20}\right)^4}\right] \wedge \left[1 - \frac{1}{1+\left(\frac{x-100}{30}\right)^6}\right] \Bigg/ x \tag{3-10}$$

- war but not too war $= war \cap \neg peace^2$

$$= \int \left[\frac{1}{1+\left(\frac{x}{20}\right)}\right] \wedge \left[1 - \left(\frac{1}{1+\left(\frac{x}{20}\right)^2}\right)^2\right] \Bigg/ x \tag{3-11}$$

- Extremely war=

$CON(CON(CON(war))) = ((war^2)^2)^2$

$$= \int_X \left[\frac{1}{1+\left(\frac{x-100}{30}\right)^6} \right]^8 \Big/ x \qquad (3\text{-}12)$$

We assume two input variables as X and Y as whine and sword, and "low", "medium" and "high" for MFs for each input. Bye assuming three themes parts "low" "medium", and "up" for sounds and three themes within the main part as "low", "medium" and " "up", and having the following rules:

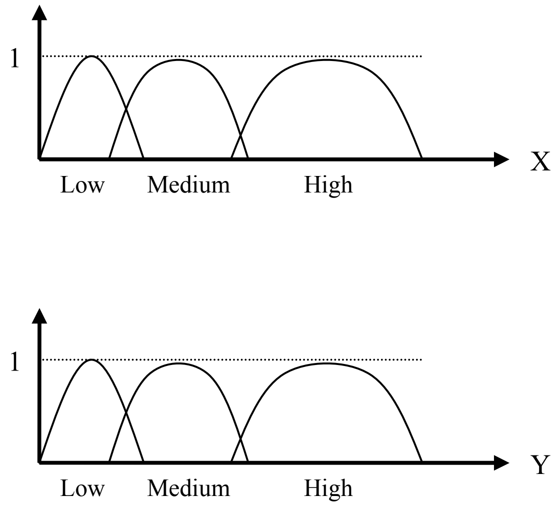

Figure 4. Fuzzy membership functions

Rule: if X is A_i and Y is B_i, then $f_i = p_i x + q_y y + r_i$;

So we will have:

Rule 1: if X is low and Y is low then LL
Rule 2: if X is low and Y is medium then ML
Rule 3: if X is low and Y is high then UL
Rule 4: if X is medium and Y is low then ML
Rule 5: if X is medium and Y is medium then MM
Rule 6: if X is medium and Y is high then UM
Rule 7: if X is high and Y is low then UL
Rule 8: if X is high and Y is medium then UM
Rule 9: if X is high and Y is high then UU

Figure five, illustrates how this two dimensional input space is partitioned into nine overlapping fuzzy regions, each of witch is governed by a fuzzy if-then rule shown above.

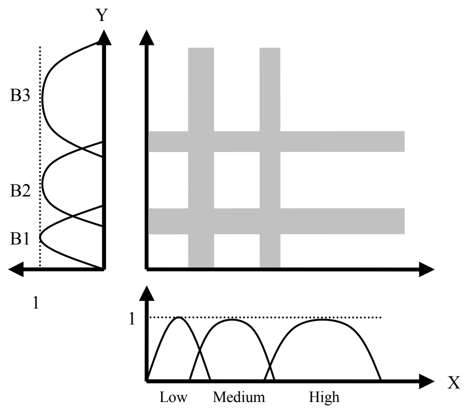

Figure 5. Input space partitioned into nine fuzzy regions

For illustration, we denote the output of the ith node in layer l as $O_{1,i}$, where:

Layer 1: Every node i in this layer is has a node function:

$$O_{1,i=\mu_{A_i}(x)}, \quad \text{for i = 1,2,3 or} \qquad (3\text{-}13)$$

$$O_{1,i=\mu_{B-3}(y)}, \quad \text{for i = 4,5,6,} \qquad (3\text{-}14)$$

where x (or y) is the input to node i from observed scene and A_i (or B_{i-3}) is a linguistic label (small, medium, high) designed via composer's believes that here we have assumed be a bell-shaped function:

$$\mu_A(x) = \frac{1}{1+\left|\frac{x-c_i}{a_i}\right|^{2b}} \qquad (9\text{-}15)$$

Layer 2: In this layer, each node output represent firing strength of a rule that is the product of the all incoming inputs:

$$O_{2,i} = \omega_i = \mu_{A_i}(x)\mu_{B_i}(y), \quad \text{i = 1,2.} \qquad (9\text{-}16)$$

Layer 3: To normalize firing strengths, the ith node in this layer calculates the ratio of the ith rules firing strength to the sum of all rules firing strengths:

$$O_{3,i} = \varpi_i = \frac{\omega}{\omega_1 + \omega_2}, \quad i = 1,2. \qquad (9\text{-}17)$$

Layer 4: Every node i in this layer is an adaptive node with a node function:

$$O_{4,i} = \varpi_i f_i = \varpi_i (p_i x + q_y y + r_i), \quad i = 1,2 \qquad (9\text{-}18)$$

where ϖ is a normalized firing strength form layer 3.

Layer 5: The input node in this layer is a fixed node labeled ∑, which computes overall output as the summation of all incoming inputs that would be final prediction of the third part of the model by composer appearing in his or her final notes:

$$O_{5,i} = \sum_i \varpi_i f_i = \frac{\sum_i \omega_i f_i}{\sum_i \omega_i} \qquad (9\text{-}19)$$

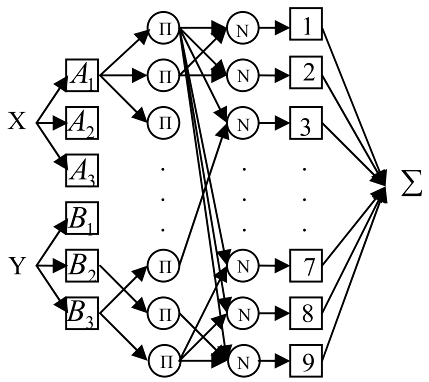

Figure 6. ANFIS architecture for a two-input Sugeno fuzzy model with nine rules

5. Conclusion

In spite of vague nature of aesthetic for describing and ruling it, we showed the ability of fuzzy logic in modeling aesthetic by using ANFIS to handling this vagueness in the process of decision making that was our aim to model. Then it was applied in designing a composer notes. Therefore, we made it more clear in this paper now the possibility of combining aesthetic and soft computing. This method, which modeled here specifically for music composing, has the ability of utilizing in many other areas as well.

6. Acknowledgements

The authors wish to thank the reviewers for their insights and helpful suggestions. The authors wish to thanks the Ministry of Science & Technology and Innovation, Malaysia and Universiti Teknologi Malaysia for supporting the research fund and conducive environments in conducting the research under the Vot 78099.

7. References

[1] Jyh-shing Roger Jang, "ANFIS: Adaptive-Network-Based-Fuzzy Inference System", IEEE transactions on systems, man, and cybernetics, May/June 1993, Vol. 23, No.3. pp 665-685.

[2] Mashinchi M.R., "On Aesthetic Cognition and Its Implementation in Computer Systems", Proceedings of the Conference on Computer Science and Information Technology, Yerevan, Armenia, 2005, pp.260-265.

[3] Mashinchi, M.R., "Implementation Of Aesthetic Cognition In Computer Systems", March 2006, Advances in Fuzzy Mathematics, Research India P.ublications, Vol. 1, No 1, pp 45-54.

[4] Mashinchi M. "Inexact Mathematics and Intelligent Systems", Computer Report, Vol. 18, No.130, pp 25-30, in Persian.

[5] Ahmadi B., "Truth and Beauty, Lectures on the Philosophy of Art", Nashr-e Markaz, Tehran, 1996, 3rd edition, in Persian.

[6] Craik, Kenneth J.W., *The Nature of Explanation*, Cambridge University Press, Cambridge,1943.

[7] J.-S. R. Jang, C.-T, Sun, E.Mizutani, , "Neuro-Fuzzy and Soft Computing", by Prentice-Hall, Inc,1997.

[8] Mashinchi M.R., Developing Models for Aesthetic

Knowledge Processing, SEUA, Yerevan, Master thesis , June 2006.

[9] Minsky Marvin, A framework for representing knowledge, Winston, 1975 211-280. Reprinted in Haugeland 1981, 95-128.

[10] Mukaidono Masao, Fuzzy logic for beginners, World Scientific Publishing , 2001.