# Using the Fuzzy Inductive Reasoning methodology to improve *coherence* in algorithmic musical beat patterns

Iván PAZ-ORTIZ [1], Àngela NEBOT, Francisco MÚGICA, Enrique ROMERO
[a] *Llenguatges i Sistemes Informátics, Soft Computing research group, Barcelona Tech.*

**Abstract.** In the present work, the Fuzzy Inductive Reasoning methodology (FIR) is used to improve *coherence* among beat patterns, structured in a musical A-B form. Patterns were generated based on a probability matrix, encoding a particular musical style, designed by experts. Then, all possible patterns were generated and the most probables were selected. A-B musical forms were created and the *coherence* of the sequence was evaluated by experts by using linguistic quantities. The output pairs (A-B pattern and its qualification) were used as inputs to train a FIR system, and the variables that produce *"coherent"* outputs and the relations among them where identified as rules. The extracted rules are discussed in the context of the musical form and from the psychological perception.

**Keywords.** Fuzzy Inductive Reasoning, musical coherence, algorithmic composition.

## Introduction

Automated algorithmic composition systems are now well-understood and documented [2, 10]. On the search for designing more effective systems with greater *expressiveness*, latest attempts have shown the need to extract representations for capturing and managing high level musical features like *coherence* or *composer personality* [5]. However, these appear commonly as a side effect of the research made in machine learning for the construction of composition or interactive systems. The fact that machine learning processes have effectively captured such features to a great degree is still object of discussion. Moreover, designed systems have not extensively incorporated perception and semantics of the generated music, including the listener psychology sensation of the musical form. Attempts to do this often deal with machine listening technics that need high computational capacity, using modules with a pre-established, symbolic domain for output's evaluation and adjustment [4]. Fuzzy systems require less amount of resources, and are not restricted to pre-established structures for the evaluation modules, allowing systems to include humans (with their psychological perspective) without having predefined representations of the desired output. In this work, we used the Fuzzy Inductive Reasoning Methodology [11] as a module to evaluate the *coherence* between two algorithmically produced beat musical patterns. This allows the system to extract the musical

[1]E-mail: Iván: ivanpaz@ciencias.unam.mx, ivanpaz@lsi.upc.edu, Ángela: angela@lsi.upc.edu

representation of the expert and translate it in terms of combinations of variables, which allow consistency between musical parts, through the subjective evaluation to listen combinations. The structure of the work is: Section 1 the basic concepts. Section 2 general methodology. Section 3 results and presents the discussion.

## 1. Basic concepts: music coherence and Fuzzy inductive reasoning methodology

In this work we explored the **musical coherence** between two patterns arranged in an A-B form. For methodological reasons we defined *coherence* as "how good A-B patterns *match* together" when they are perceived by a listener. The evaluation was made by using linguistic variables [9]. The coherence will depend on the contrast and repetition points between A and B and on the moments on which those are situated.
A complete documentation of FIR can be found in [6,1]. The system is fed with raw data from the system under study. It has four basic functions: The *fuzzyfication* process transforms the data into triplet format. The qualitative modeling utilizes a *fuzzy optimal mask function*, referring to the selection of which variables participate in the output prediction. This process finds the qualitative relationships between the different input variables. This analysis is performed by using either an exhaustive search or by means of search trees or genetic algorithms. FIR uses Masks as qualitative models, by analyzing the episodical behavior (recorded in the FIR data matrix) of the system for the identification of a qualitative modeling used for future forecasting. FIR creates the best mask for only one input variable, the best for two, and so on. The masks are called of complexity one, two, etc. The quality of the mask (Q) is determined by the uncertainty reduction measure based on the entropy associated with the transition matrix of states associated with the set of variables of the mask. The *fuzzy simulation* process allows the model to predict future qualitative outputs based on past experiences by interpolation processes in the input variables to extrapolate the output. The *regeneration* module performs the inverse process of *fuzzyfication* by transforming the triplet into the original data format.

## 2. System design and methodology

The beat patterns were created in the context of UK garage/two step [13] for 3 instruments: *kick*, *snare* and *hihat* based on the analysis of [2]. The probability vectors below describe the independent probability, in the interval [0,1], for each instrument to play in a particular moment. Each vector represents a 4/4 bar where each quarter (1 unit) is divided in four sixteenth notes (1/16 of unit).

[0.7, 0.0, 0.0, 0.1, 0.0, 0.0, 0.0, 0.2, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.3] kick

[0.0, 0.0, 0.0, 0.0, 1.0, 0.0, 0.0, 0.5, 0.0, 0.2, 0.0, 0.0, 1.0, 0.0, 0.0, 0.3] snare

[0.0, 0.0, 1.0, 0.0 , 0.0, 0.0, 1.0, 0.7, 0.0, 0.0, 1.0, 0.0, 0.0, 0.0, 1.0, 0.7] hihat

To avoid the cases where more than one instrument play at the same time, we considered three (musical hierarchy) rules: If kick and snare, then kick. If kick and hihat, then kick. If snare and hihat, then snare. From all possible patterns we selected the 10 with highest probabilities which yielded a set of 20. Those were sequenced in A-B form and reproduced to the listener at 120 beats per minute, in sequences of 4 times A followed by 4 times B, for psychological perception reasons. The coherence between A-B patterns was evaluated using linguistic variables: low, medium or high. We considered 105 different A-B forms. The data, was structured in the format: [1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22,23,24,25,26,27,28,29,30,31,32,33] Where numbered variables correspond with each one of the 32 sixteenth in the A-B pattern. Entrance 33 is the listener's evaluation of the sequence, and is the output of the system. The values of the entrances of the vector were defined as 4- non strike, 1- kick, 2- snare, 3- hihat. The output took *coherence* values of 1: low, 2: medium, and 3: high. With these considerations we fed the FIR, setting the membership values at the center of the bell membership functions in the *fuzzyfication* process to allowed the model to manage the crisp data. To find the relations among variables we used a comprehensive search based on Shannon entropy, i.e searching for the set of variables that make the state-transition matrix as deterministic as possible. Those are the relevant variables. We also used the Linguistic Rules in FIR algorithm (LR_FIR, [1]), which is a rule-extraction algorithm, that starts from the set of pattern rules obtained by the FIR model previously synthesized, and is able to derive linguistic rules from it.

## 3. Results and discussion

We were interested in modeling the different configurations from the relevant variables that produce a coherent perception in the listener. These are consistent with the musical structure of the style. The results allow us to change the rhythmic motives so that a new system produces parts perceived as more coherent. Extracted rules showed that either V4 nor V20 variate. This should be attributed to the fact that the probability of having a strike in these variables is determined by the kick with probability of 0.1. This result left only the variables: 1, 8, 10, and 16 creating variation in part A, and 17, 24, 26, and 32 in part B. The extracted rules using the LR_FIR in cases 1) considering the 8 variables and 2) considering the most relevant variables obtained by FIR, which correspond to the variables 17, 24 and 32, are displayed in Figure 1 Left and Right, respectively. The rules describe which is the value (instrument) of a particular variable (e.g V17-1 should be read: "variable 17 was in 1 (kick)"), and how the combination of variables produces a particular output. The third rule of Figure 1 should be read as: IF V16 IS 1 AND V26 IS 2 AND V32 IS 1 THEN V33 IS 3.

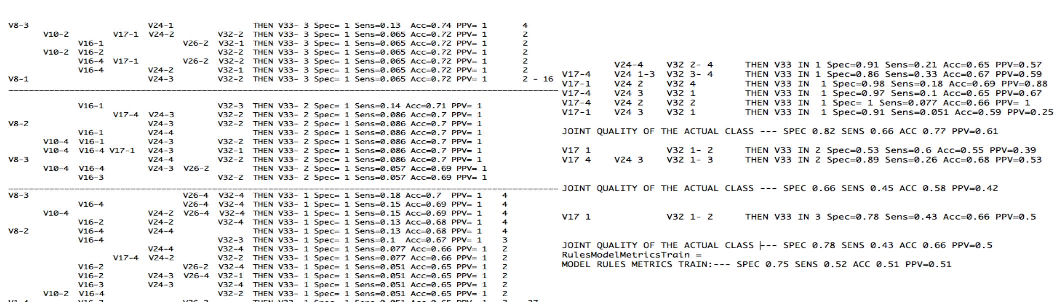


**Figure 1.** Extracted rules for the input data by using LR_FIR considering 8 input variables and one output.

In left down, are the rules for evaluated low coherent patterns *LC* (i.e. the output of V 33 is 1). We found a great amount of silences (Variables in 4), specially in V32. If we compare this behavior with rules for medium coherent *MC* patterns (center of the figure) and with high coherent *HC* patterns (top) for the same variable, we can say in general, that silences in V32 affect the coherence of the perceived sequence. An explanation for this is given in terms of the subjective perception of patterns. Given that the patterns are composed by 4 times A followed by 4 times B, the sensation of periodicity of B will be produced in great amount by the variable that completes the cycle by connecting it with their repetition, which is V32, in this case connected with variable 17, from which we know (from the probability matrix) that 70% of the cases will be 1 (kick). As we said, this behavior is expected from the point of view of the subjective perception of rhythm. In this case, if pattern "A" has been interesting enough, the focus in searching coherence will be on pattern B. And to perceived B as cyclically coherence we need to look at V 32. The previous hypothesis can be also supported by the rules: V10-2 V17-1 V24-2 V32-2 THEN V33- 3 and V16-4 V17-1 V26-2 V32-2 THEN V33- 3. Belonging to patterns evaluated as *HC*. In both cases, V17-1 and Vd32-2. These two rules represent approximately 4/16 of the cases evaluated as *HC*. Moreover, if we look into the different masks (Table1) when we looked for the mask for one input variable, we obtained V32. Also, sequences evaluated as *MC* (Left center), do not have silences in V32. In the right are the extracted rules considering three Variables. We can see that we only have one rule describing the *HC* evaluating cases: V17-1 V32 1- 2 THEN V33 IN 3. In which the behavior described above is clearly expressed. V17 most be 1, and V32 could be 1 or 2. As said, those variables determine the cycle sensation in B. Also all silences, with one exception, are found in patterns evaluated as *LC*. The different masks created by LR_FIR for one to eight variables are shown in Table 1. They represent the variables that

**Table 1.** Masks created by LR_FIR for one to eight variables. The quality of the masks is denoted by Q

| V1 | V8 | V10 | V16 | V17 | V24 | V26 | V32 | Q |
|---|---|---|---|---|---|---|---|---|
| - | - | - | - | - | - | - | * | 0.14 - Q1 |
| - | - | - | * | - | - | - | * | 0.27 - Q2 |
| - | - | - | - | * | * | | * | 0.30 - Q3 |
| - | - | - | - | * | * | * | * | 0.28 - Q4 |
| * | * | * | * | * | - | - | - | 0.27 - Q5 |
| * | * | * | * | * | - | * | - | 0.21 - Q6 |
| * | * | * | * | * | * | * | - | 0.22 - Q7 |
| * | * | * | * | * | * | * | * | 0.20 - Q8 |

have more influence in the prediction of the output [1]. At the top the mask for only a single variable "Q1" contains the V32. As discussed, we can explain this by considering this variable as the one who, together with V17 (which is 1 in 70% of the cases), give a cyclic sensation to the pattern B. Q2 is in terms of V16 and V32. This selection can be explained considering the role of V32, and that V16 plays for A-patterns the same role of V32 for B-patterns, so we can understand this in terms of the cyclic sensation they produce. Also, V16 is a connection variable between parts A and B of the pattern, so it gives V16 another important role in the perception of the whole. In the case of Q3, the selected variables were 17, 24 and 32 which are related with the cyclic perception of the B-pattern (17 and 32). V24 is playing a role in increasing the rhythmic interest. The same idea explains Q4, where 24 and 26 were selected. However, an interesting behavior appears in Q5, when, with exception of V17, all selected variables belong to pattern A. The following masks add 26, 24 and 32, respectively. This is explained because when new variables are added to the mask, the new variables together, can explain a great amount of the overall perception in comparison with the original ones.

## References


[1] Castro, F., Nebot, A. and Mugica, F., 2011. On the extraction of decision support rules from fuzzy predictive models. Applied Soft Computing, 11 (4), 3463-3475.

[2] Collins, N., 2003 Algorithmic Composition Methods for Breakbeat Science ARiADA No.3 May 2003.

[3] Collins, N., 2009 Musical Form and Algorithmic Composition. Contemporary Music Review Vol. 28, No. 1, February 2009, pp. 103-114

[4] Collins, N., 2012. Automatic Composition of Electroacoustic Art Music Utilizing Machine Listening. Computer Music Journal, 36:3, pp. 8Ð23, Fall 2012 Massachusetts Institute of Technology.

[5] Duncan, W., Kirke, A., Miranda, E,. Etienne B., Roesch, Slawomir J. Nasuto. Towards Affective Algorithmic Composition. Proceedings of the 3rd International Conference on Music & Emotion (ICME3), Jyväskylä, Finland, 11th - 15th June 2013. Geoff Luck & Olivier Brabant (Eds.)

[6] Escobet, A., Nebot, A., Cellier, F. E., 2008. Visual-FIR: A tool for model identification and prediction of dynamical complex systems, Simulation Modeling Practice and Theory, vol. 16, nij 1, pp. 76-92.

[7] Hannon, E. E., Snyder, J. S., Eerola, T., & Krumhansl, C. L. 2004. The Role of Melodic and Temporal Cues in Perceiving Musical Meter. J. of Experimental Psychology: Human Perception and Performance, 30(5), 956Ð974.

[8] Kirke, A., Miranda, E. R., & Nasuto, S. (2012). Learning to Make Feelings: Expressive Performance as a part of a machine learning tool for sound- based emotion therapy and control. In Cross- Disciplinary Perspectives on Expressive Performance Workshop. Presented at the 9th Int'l Symp on Computer Music Modeling and Retrieval, London.

[9] Kosko B (1986) Fuzzy Cognitive Maps. Man-Machine Studies, 24, 65-75.

[10] Nierhaus, G., 2009. Algorithmic Composition. Paradigms of Automated Music Generation. Springer Wien, New York.

[11] Nebot, A., Mugica, F., 2012. Fuzzy Inductive Reasoning: a consolidated approach to data-driven construction of complex dynamical systems. International Journal of General Systems, 41(7), pp. 645-665

[12] Pareyón, G., 2011. On Musical Self-Similarity Intersemiosis as Synecdoche and Analogy. International Semiotics Institute at Imatra Semiotic Society of Finland. Printed by Yliopistopaino, Helsinki.

[13] Shapiro, P. 1999. Drum 'n' bass, the Rough Guide. Rough Guides Ltd.

[14] Spring, G., and Jere T. Hutcheson. 1995 Musical Form and Analysis. Madison, Wis: Brown & Benchmark, Print.

[15] Typke, R., Wiering, F., Remco C. Veltkamp. 2005 A survey of Music Information Retrieval Systems. Queen Mary, University of London.

[16] Zadeh, L., Fuzzy Sets. 1965 Information and Control 8, 338-353.