A Model of Pattern Processing for Music
Haruhiro Katayose and Seji Inokuchi
Laboratories of Image Information Science and Technology
Tel: +81-6-873-2030 / E-mail: katayose@image-lab.or.jp

ABSTRACT: This paper describes a model of pattern processing for music, based on the cost of remembrance. The algorithm works to make a hierarchical pattern model with the lowest remembrance cost, which is defined by the function of pattern definition cost, the distance between the model and the pattern instance, and the number of instances.

1. Introduction
To induce patterns in music and to predict the next sequences are primal musical skills. These skills are important both offline musical processing such as interpretation and in realtime musical processing. Many researchers have been studying pattern induction and developed matching procedures, some of which have been used for interactive music systems (Witten, L.; Rowe, R.). One of the biggest problems of realizing pattern induction in music is how to define similarity of patterns. It seems to me previous studies have not paid as much attention to structural stability, as the criterion for judging sequences belonging to a class that is the same or similar. Here, structural stability functions as a means to gauge how repetition of same or similar patterns may weaken the relative criterion for judging musical sequences belonging to a similar class. The aim of this study is to make an integrated model which considers the structural stability and the thresholds for the similarity. It is an attempt to handle "remembrance capability."

2. Conceptual Basis of Pattern Processing
There is a concept called AIC (an information criterion) which was proposed in the field of statistics. AIC is defined as \((-2) \ast \log (\text{likelihood}) + 2 \ast \text{(number of parameters)}\). The first term takes charge of how much the model fits the signal. The second term works to choose the model with a lesser number of parameters. One of the most interesting points of AIC is to consider how easy the signal is remembered. The idea of the pattern processing method proposed here is based upon this "principle of parsimony": economization of memory. It would work to take the hierarchical pattern model with the lowest remembrance cost, which is defined by the function of the pattern definition cost, the distance between the model and the pattern instance, and the number of instances. Using these parameters, the phenomenon that, "unlearned pattern structure varies when new event comes", is modeled.

2.1 Pattern Definition Cost (PD)
This cost is responsible for the information quantity of the model (base) of a pattern. PD is given by the sum of the total number of the parameters to describe the all events in the model and the log-order likelihood related with each parameter. The pattern is essentially multi-dimensional. Some of the parameters are note-name, surface-rhythm (note_onset), prevailed-chords, and so on. Note-names may be reduced into pitch differences between adjacent notes and a constant. As for rhythm, quantized beat and delicate time can be candidates of the dimension. Information quantity depends on the circumstances such as style or subjects' experience. But it is possible to make an experiment on a given circumstance.

2.2 Distance between Model and Instances (DMI)
DMI corresponds to the likelihood of the patterns for belonging to a class. DMI is given by the total log-order deviation between the model and the established instance for given parameters.

J O U R N A L O F C O M P U T E R S I C E N C E 1 9 9 5 5 0 5
2. Minimization of Total Cost.
The cost of a music sequence is defined by the following equation: \( \alpha \times (PD + DMI) + NI \), where NI is the cost regarding number of pattern instances and \( \alpha \) is weight ratio of likelihood tolerance / structural stability. The basic idea is to search the model of the pattern and instance, as the total cost might be minimum at the given \( \alpha \). Another important point of the idea is the hierarchical utilization of the model pattern to define larger model pattern. It also contributes the total cost reduction and reflects "easy to remember patterns." The cost regarding this sub-model-pattern placement is calculated from the number of existing smaller model patterns and the information of the positional identifier.

3. Pattern Processing Procedure
3.1 Repetition Interval
The simplest way to find patterns is to implement the round search. Actually the method adopted here is a round search in a broad sense. The computational cost of the round search is extremely high and seems to be far from human mind procedure. In order to reduce search calculation, the proposed procedure utilizes repetition interval and pattern initiating candidates (Katayone, H.). The repetition interval is detected using the auto-correlation method for the each vector of pitch differences, contour and note-onset.
Some dominant repetition intervals are used for the following procedure. In case this procedure is applied to real-time processing, ten seconds or so is required for the appearance of fundamental repetition.

3.2 Pattern Initiating Candidates
What is the starting point of the patterns or groups? It is a significant problem in cognitive science. The candidates of pattern initiating are the very starting point of the event sequence and the structural accent points supported by gestalt. The structural accent consist of the events after rests, the big contour accents, and durational accents. It is difficult to pick up the definite structural accent points, but they can be used in order to avoid redundant search. In other words the true structural points are detected when the algorithm finds the pattern with the minimum cost.

3.3 Search Based on Repetition and Pattern Initiating Candidates
The objective of search is to find the hierarchical pattern with the minimum remembrance cost. The first task is to find the model pattern. The program tries to find if there exists the same or similar pattern initiates on a pattern initiating candidates and on the points failing on the integer times of the repetition intervals. If more than two similar patterns are detected, the program registers the pattern as the model with the PD cost. There is the possibility of plural interpretation when overlapping groups are employed. At this point, groups inconsistent with each other or of bigger cost are maintained. When the program analyzes the upper layers and assigns the stored groups, there by fixing the pattern model instance, the contradicting overlapping models are forgotten, while other groups are maintained.

4. Summery
Due to page limitation, this paper has been focusing on the conceptual basis of a pattern processing model for music and the prival procedure for the implementation. The experimental results and the discussion about the real-time application will be shows in the next paper.

References