Institute of Acoustics

A MICROCOMPUTER BASED SYSTEM FOR EVOLVING MUSICAL STRUCTURES

Jean-Philippe Gervais and Michael Greenhough

Physics Department, University College, Cardiff

INTRODUCTION

During recent years, the music industry has seen a lot of computer based tools being developed to help musicians at various stages of their work. Such tools include improvisation and performance systems, powerful sequencers, musical notation editors, and sound synthesizers. A great effort has been made to improve the user interface of these new pieces of equipment, by introducing menu-driven systems, windowing techniques and screen icons, and new control devices such as the joystick, mouse, and drawing pad.

However, the systems for which these facilities have been developed, often require from the user a detailed knowledge of the way parameters controlled (or generated) by the system are operated upon.

In this paper, we propose a new approach to be used to control, in real-time, a system which generates musical structures from a stochastic algorithm. The approach originates from an optimizing technique known as Evolution Strategy [1] and takes its roots from natural evolution.

It will be shown that by monitoring the user's assessment of its output, the system can automatically modify some of its parameters, in order to gradually converge towards satisfactory material. Before detailing the new method, a brief description of the system and its present means of control will be given.

THE BASIC SYSTEM

System Description

The system for which the new method of control is being developed, is microcomputer based and can generate polyphonic musical structures for 16 voices in the form of MIDI parameters (Musical Instrument Digital Interface -a hardware and software convention for communication between synthesisers and computers) which are then used to drive any MIDI synthesizer.

Parameters under control are the pitch, duration, volume and sustain of a note. The system is extensively described in [2].

At present each parameter is determined by 3 levels of control:

(i) An "a priori" level representing the space of possible values for a parameter with their probability of occurrence. Control from this level alone is exerted independently of the context.

(ii) A "1st-difference" level defining the probability of various changes of parameter values with regard to chosen context reference points.

(iii) A "2nd-difference" level controlling the changes of changes of parameter values. This can be used to shape the longer-term properties of the output, for example by controlling the curvatures of a melody.

When combined, these 3 levels of control form the constraints imposed on a random generator, necessary to determine which of the possible events is to be produced. Although more levels of control could be introduced, this three-level structure provides quite effective control, in real time, of the pitch, duration, volume, and amount of sustain (legato/staccato) of each event.

By setting appropriately the various distributions as well as the context reference points, a great variety of musical effects can be easily achieved, including:

-parallel and contrary motion between voices, canonical effects and ostinati
-the establishment of particular rhythm and syncopation effects
-accentuation,crescendo,decrescendo
-staccato,legato.

Furthermore, it is thought that timbre could be controlled by the system in a similar fashion, if it was expressed, for example, in terms of Frequency Modulation parameters.

Present system control
One direct way of controlling the system is to modify the form of any of the probability distributions at run-time. These can be displayed in groups of four, and can be altered by means of a movable cursor. In addition, single parameters read from joystick positions can be used to modify either the overall shape of a distribution or one of its particular elements. Each of the voices can be displayed in real-time in standard music notation. Other graphic facilities can inform the user about various settings such as the volume of a voice, its pitch range, and its status (on/off). Also available to the user in a form of display is information about the horizontal and vertical relationships existing between voices. All the parameters cited above can be modified in real time by single cursor movements.

A MICROCOMPUTER BASED SYSTEM FOR EVOLVING MUSICAL STRUCTURES

Limitations of present control

The often tight constraints imposed by the various probability distributions ensure that the available musical-structure "space" is highly restricted and biased towards the more intelligible and appealing kind of output. However, it can still seem a vast region if one is searching blindly for a particular musical effect. We could improve matters by simply demanding from the user a more explicit description of the musical gaol. This is however not consistent with our system-development philosophy which is to exploit the passive abilities of users. Furthermore, asking for directions to a specific goal will tend to preclude some new and interesting regions of musical-structure space which more open-minded wandering might lead to.

EVOLUTION STRATEGY - AN ALTERNATIVE MEANS OF CONTROL

The problem of arriving at structures which are satisfactory for some purpose without the need to describe them explicitly suggests an evolutionary strategy. Here the ear-brain system of the user constitutes the environment and the successively refined samples of musical output correspond to individual organisms of an evolving species. The process of selection corresponds to the user signalling approval or disapproval to the system which then reacts by appropriately adjusting the probability of re-occurrence of features in the favoured or offending sample. Thus the probability distributions which control the musical output constitute a kind of " gene-pool" which becomes enriched with some features and depleted of others as user-system interaction progresses. This analogy is explored further in [3].

In order to be able to control the probability of re-occurrence of features in the sample, a feedback mechanism is used, whereby information present in the sample is analysed and used to modify appropriately the probability distributions which gave rise to it.

Feedback mechanism

If we aim to "steer" a system through some kind of feedback mechanism a comparison will often be made between the expected and observed values taken by each parameter. In the system described here, any sample taken from the output reflects the probability distributions which generated it, because of the close relationship existing between the probability of occurrence for an event and its observed frequency, .

For example, let P represent an element of the initial "a priori" distribution, A the frequency count of this element in the sample, and F some weighting factor. We define P^*, the corresponding element of the modified "a priori" distribution, as

$$P^* = P\left(\frac{A}{P}\right)^{\pm F} \qquad (1)$$

In the modified "a priori" distribution, then, each element is scaled by an amount proportional to the weighted ratio of its frequency in the sample over its probability in the initial "a priori" distribution.

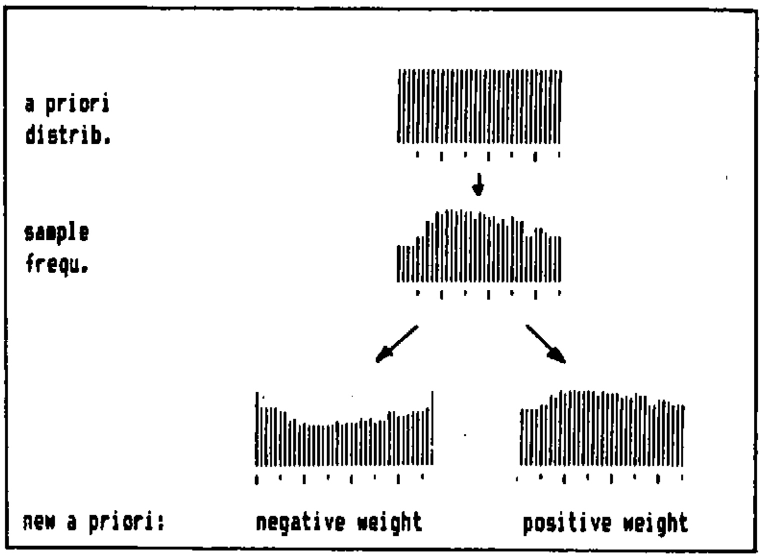

Fig.1. Illustrating the effects of positive and negative weightings

Description of experimental principles

To evaluate objectively the effectiveness of the feedback mechanism described above, we have devised an experiment, in which users have to steer the system in a desired direction merely by indicating how closely the musical samples reflect some given target attribute. The first task chosen for these experiments, was to force the system to play notes belonging to a restricted region of its "a priori" pitch distribution. The user is given a reference note representing the centre of the region to be selected by the system, and then a sample of fixed length is generated and played.

After each sample, the user has to decide if the sample was very far, far, not so far, close or very close with respect to the target area. Once a decision is reached, the system computes from equation (1) a new "a priori" distribution which will be used to generate the next sample. This process is then repeated until the user feels the task has been completed satisfactorily.

We report here some initial experiments performed to determine what approximate ranges of various parameters would lead to a convergence of the search in a reasonable time. As a preliminary we briefly consider some of these parameters to develop a "feel" for their likely effects on the search.

The size of the output sample

Large samples, if clearly judged as very good or very bad, will help accelerate the convergence to the target. We can illustrate this by the possibly absurd extreme case where a sample is output which is totally satisfactory and of the full appropriate length for the purpose. The user need only pause to "save" the sample and quit the experiment. However, because of the stochastic nature of the generation process, the larger the sample the smaller will be the chances of a clear-cut success or failure.

Small samples will be much easier to judge as appropriate or inappropriate as long as the criteria are fairly simple. However, the perturbation of the probability distributions resulting from a judgement will be correspondingly small and we might expect the speed of convergence to suffer. Possibly the size of the sample could be increased as it becomes more satisfactory. However, we might expect the optimum size to be located towards the smaller end of the range. After all, the evolution of complex life forms within the lifetime of the earth is only explicable if it consisted of a large number of small steps, each of which is "tested" by the environment.

The number of possible user responses The minimum would be any two chosen from "good", "bad" and "indifferent". This extreme requires minimal discrimination and consistency from the user. A finer range of responses, for example a 7-point scale, would provide more information per trial but would be proportionately more demanding of the user.

The weightings attached to responses Large weightings will speed the convergence but may lead to instability, for example, through an oscillation of the search. Also, through the consequent coarsening of the search, some possible solutions may be missed. Smaller weightings will result in a slower but more thorough and reliable search.

It should be clear that these various parameters would not exert their influences on the course of the experiment in anything like an independent fashion.

First experiments

Some early experiments have shown that for a simple task such as "persuade the system to produce notes belonging to a restricted pitch region ", it is possible to converge on the target after a relatively small number of samples, when their length is of about 8 to 16 notes. Presently, the user has the choice between 5 options to judge the qualities of each sample and the weighting factors used vary between 0 and ± 1.

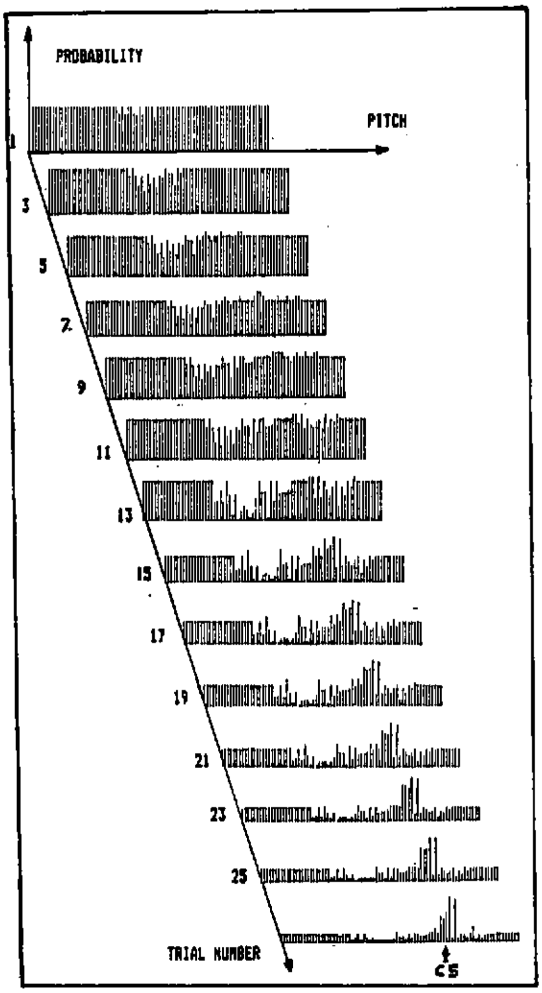

Fig.2 shows the evolution of an "a priori" distribution starting from a equiprobable chromatic distribution and trying to select notes around pitch C5.

length of sample:8 notes
target: notes around C5

responses:
- very close =>F := 0.02
- close =>F := 0.01
- neutral =>F := 0
- far =>F := -0.01
- very far =>F := -0.02

<u>Influence of the weight on the convergence</u>

In order to analyse the influence of the weighting on the rate of convergence we have conducted another experiment in which the system produces a sample and analyses how well it fits the frequency distribution expected from a chosen, target "a priori" distribution. The system then modifies the original "a priori" distribution so that it will produce a sample more representative of the target frequency distribution. The experiment terminates when the two distributions differ by less than some chosen "error" magnitude.

This method is purely objective as it does not involve the responses of a user. The decision taken after each sample results from a binary test. Because of the nature of the test this experiment uses a restricted number of notes. The weightings used have been varied from 0.08 to 1 and results indicate that the greater the weight the fewer the number of samples is needed before convergence. The number of trials before convergence for each weighting are plotted in Figure 3. Convergence tests have been repeated 15 times for each separate weighting.

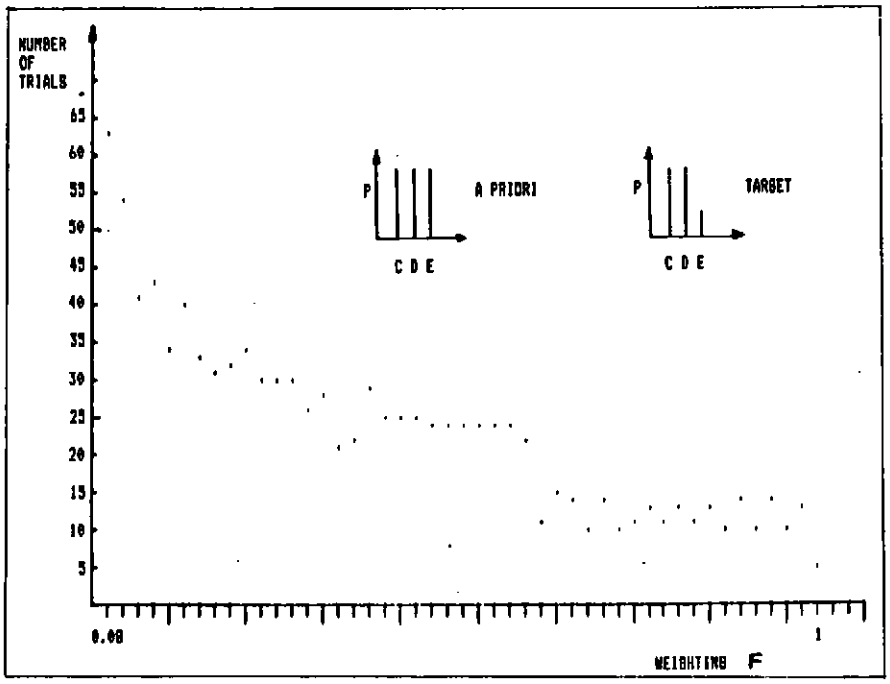

Fig.3. Number of trials to convergence as a function of weighting factor.

CONCLUSIONS

Although only a few experiments have been carried out so far, the results obtained seem to indicate that the feedback mechanism in use is capable of causing the shape of a probability distribution to tend towards a favoured one. Many more tests need to be done to determine the optimal number of notes needed in each sample, the optimal weighting to be used and the best number of user responses. Of further importance is to find out how these parameters interact with each other.

To be consistent with the present philosophy of the system, the whole process of generating samples, analysing them and modifying the distributions will have to be made a continuous rather than quantised one. Further work will include the application of this control technique to other levels of pitch control in the system. It is hoped that the particular tonality and the shape of a melody could be evolved in a similar fashion. If the technique proves to be satisfactory for controlling all the aspects of pitch, investigations will be made to see how we can best apply the same method to control other parameters such as duration and volume.

A MICROCOMPUTER BASED SYSTEM FOR EVOLVING MUSICAL STRUCTURES

ACKNOWLEDGEMENTS

J-Ph Gervais is grateful to The British Council for supporting this research.

BIBLIOGRAPHY

[1] Rechenberg, Ingo.
Evolutionsstrategie (Frommann Verlag, Stuttgart: 1973)

[2] Greenhough M.
A stochastic algorithm for the real-time control of musical structures
Proceedings of the Institute of Acoustics, Vol 7 Part 3 (1985). 157-164.

[3] Greenhough M.
The Philosophy and Design of a Real-Time Music-Structure Generating System
Interface - Journal of New Music Research (The Netherlands) (In Print)