AUTOMATIC PERFORMANCE OF MUSICAL SCORES BY MEANS OF NEURAL NETWORKS: EVALUATION WITH LISTENING TESTS.

G. U. Battel*, R. Bresin+, G. De Poli+, A. Vidolin*

*Conservatorio di Venezia "Benedetto Marcello"

+D.E.I. - C.S.C. Università degli Studi di Padova
via Gradenigo 6/a - 35129 Padova - Italy
phone: +39 49 8287631- fax: +39 49 8287699
E-mail: depoli@dei.unipd.it, rb@csc.unipd.it

Introduction

Musicians, according to the instrument they play, make loudness, duration and timbre deviations on the notes of the score they are performing, since the traditional musical notation does not suffice the composer's real intentions, and leaves some freedom's degrees to the player himself. These deviations determine the performing characteristics of a pianist in respect to another one. Furthermore a literal performance of a musical score would lead to an extremely mechanic and unnatural performance to the listener's ears.

The present work starts from the Sundberg's and co-workers' researches on automatic scores' performance [1] [2] and continues the research on real-time piano scores performance by mean of particular artificial neural networks. In our previous works [3] [4] we showed the possibility to build some neural networks which can learn some performing rules. These nets show good generalization properties, and, after a training phase, are able to do real-time performances of any score introducing some appropriate deviations.

In the present research we propose a comparison test between various performances to evaluate, by mean of listening tests, the use of trained neural nets in automatic performance.

Description of the experiment

Methods

Materials

Two tonal melodies were chosen to be used in this study.

The first one is the theme of the third tempo of the Mozart piano sonata K. 284, the second is the theme of the first tempo of the Mozart piano sonata K. 331.

For the experiment it has been used a subset of Sundberg's rules: the selected rules influence mainly the relations between near notes and don't involve greater segments. In this way the selected rules are more suitable to a structurally simple musical example and to a strictly classical repertoire. This formal need, and since the sonata K. 331 was previously used in other works [1], brought us to the choice of the two themes of Mozart. The starting hypotesis are two: it is possible to obtain an acceptable performance with small performing deviations and without involving the great form; for the intrinsic characteristics of Mozart's music, the melodies are performable in a meaningful way on their own, and the deviations sound pleasent and understandable to the listener.

We performed each of the two melodies in three ways: deadpan (with no expression); with expression given by a subset of Sundberg's performance rules, in the following these melodies will be called rules-melodies; with expression given by two neural networks trained with the same subset of Sundberg's rules, in the following these melodies will be called nn-melodies. The rules that we applied are the following [1]:

- durational contrast
- melodic charge
- articulation of repetition
- leap tone duration
- leap articulation
- high loud
- phrase.

As an example of comparison, in figure 1 are showed the time (in milli-seconds) deviations in the two non deadpan performances of the theme of the K. 284 sonata. The 0 value is the nominal value, it corriponds to the deadpan version (i.e. no time deviations).

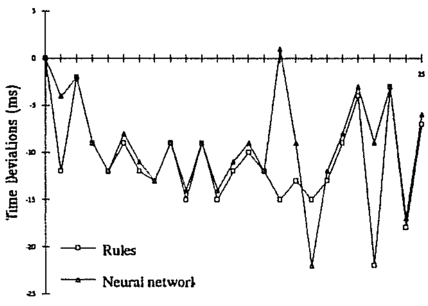

Fig. 1. Time deviations in the K 284 sonata (particular).

Equipment

The melodies were performed via MIDI by a Yamaha Disklavier Grand Piano connected to a 80386/40MHz PC compatible. To obtain the deadpan melodies and the rules-melodies we used a program called MELODIA developed at the C.S.C. [5]. To obtain the nn-melodies we trained two neural networks: one for the loudness deviations, and another for the time deviations (see Figure 2).

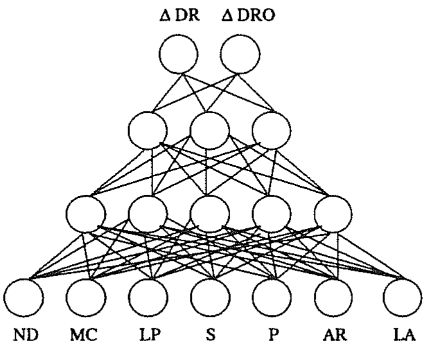

Fig. 2. Neural network for time deviations (ND= Nominal Duration; MC = Melodic Charge; LP = Leap Presence; S = Semitones in a leap; P = Phrase; AR = Articulation of Repetition; LA = Leap Articulation).

Subjects

Subjects for the study were 20 professional musicians, and students of the last years of the music conservatory of Venice, who volunteered for the experiment. They were 12 men and 8 women. The youngest was 15 years old and the oldest 32 years old. 17 of them were pianists: undergraduate and postgraduate.

Procedure

Subjects were asked before the experiment began that they had to read a paper with the instructions for the experiment and the cells where to write their judgements. Each subject was given a copy of the paper. In the paper was explained that the aim of the experiment was to compare three different piano performances built with the help of computer. In the paper were listed the title and the authors of the melodies. The text of the paper was the following:

"You will listen to three different performances of each melody, and you'll have to evaluate the musical quality of each performance as if a student is playing. You must your evaluation with a note from 1 to 10, using all the scale if possible. 1 is the worst note, 10 is the best one: avoid to give to much notes in the intermediate range (between 5 and 6), try to use extreme values (1 and 10).

The judgement doesn't have to be to critic in a absolute sense, but has to show the qualitative differences, which you find in the three different performances of the same melody. There aren't right or wrong answers: the aim of this test is to find the performance you think is the best. Between two melodies you have 30 seconds of time to judge the previous three different performances of the melody just heard, and to write it in the apposite cells."

The three different performances of each melody were played in a random order.

The total duration of the test was 4'50".

Results

The results can be analyzed by using an ANOVA with repeated measures on each of two factors (version, melody). The analysis shows significant effects for version [F(2,38)=10.36; p= .000].

The nn-melody is the most preferred version (preference rating, 6.73), followed by the rules-melody (6.3) and the deadpan (4.28).

The most important result is the preference given by the subjects to the performed melodies (rules-melodies, and nn-melodies), that obtained a mean score 2 points greater than the deadpan version. Furtheremore, if we consider the scolastic italian tradition, the subjects gave a more than fair rating to the performed versions and an unsatisfactory rating to the deadpan version, even if these values don't interest the extreme values of the scale (from 1 to 10) (see Figure 3).

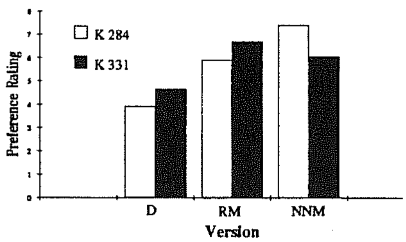

Fig. 3. Means for the interaction between the effects of version and melody (D = deadpan; RM = rules-melody; NNM = nn-melody).

Discussion

The subjects found that the greatest difficulty was the small difference between the three performances of the same melody, often they said: "They are quite equal". This fact valids the initial hypotesis, and stresses the need to continue our research considering a larger number of performing rules and a larger repertoire.

The principal outcomes of this experiment are the equivalence of the nn-melodies and the rules-melodies, and the preference they had in respect to the deadpan melodies. The reasons because the differences are not so marked are mainly two: we used k=1 in the Sundberg rules (this means a little emphasis in the performance rules, since all the melodies don't have a slow metronome); from the discussion we had with the subjects after the test emerged the difficulty to find great differences between the three performances of the same melody. The mean of the subjects find the performances to be poor, maybe for the absence of an accompaniment.

Conclusions

In our opinion, the main reason for the preference of the nn-melodies in respect to the rules-melodies is that the deviations depends from the contribution of more rules. When only one of these contributions is the responsable of the deviations, then neural networks and Sundberg's rules give the same results. When more factors act together, then the additive action of Sundberg's rules system, and the properties of interpolation proper to the neural networks give different results. From this observation and from the results of the test it comes out that neural nets follow strategies which are closer to the performing action of a human performer, and so they can simulate in a better way the process of performance. (see Figure 1)

References

[1] Friberg A. "*Generative Rules for Music Performance: A Formal Description of a Rule System*", Computer Music Journal, vol. 15, No. 2, pp. 56-71, Summer 1991

[2] Sundberg J. et al. "*Performance Rules for Computer-Controlled Contemporary Keyboard Music*", Computer Music Journal, MIT Press, vol. 15, No. 2, pp. 49-55, Summer 1991

[3] Bresin R., G. De Poli, A. Vidolin "*Un approccio connessionistico per il controllo dei parametri nell'esecuzione musicale*", Atti IX Colloquio di Informatica Musicale, Genova, pp. 88-102, 1991

[4] Bresin R., De Poli G., Vidolin A. "*Symbolic and sub-symbolic rules system for real time score performance*", Proceedings of the 1992 ICMC, International Computer Music Association, San Francisco, 1992

[5] Bresin R. *"MELODIA: a program for performance rules testing, teaching, and piano scores performing",* elsewhere in these proceedings, 1993