

Evolving Expressive Music Performance through Interaction of Artificial Agent Performers

Qijun Zhang, Eduardo R. Miranda

Interdisciplinary Centre for Computer Music Research (ICCMR)

University of Plymouth

Drake Circus, Plymouth PL4 8AA, UK

{qijun.zhang, eduardo.miranda}@plymouth.ac.uk

Abstract. We propose a model of expressive music performance (EMP), focusing on the emergence of EMP under social pressure, including social interaction and generational inheritance. Previously, we have reported a system to evolve EMP using Genetic Algorithm, exploring the effect of generational inheritance. This paper presents a system that evolves expressive performance profiles through social interaction, with a society of artificial agent performers. Each performer owns a hierarchical pulse set (i.e., hierarchical duration vs. amplitude matrices), representing a performance profile for a given piece. An agent performer evaluates a performance profile with a set of rules derived from the structure of the piece in question, and imitates others' performances if appropriate. Then it modifies its pulse set accordingly. We demonstrate that suitable performance profiles emerge from social interactions where the diversity and the commonality of evolved performances are observed in the society of agents.

Keywords: Expressive music performance, pulse set, musical structure, social learning, emergent behaviour, learning by imitation

1 Introduction

In the context of Western tonal music, music expressions are delivered in a performance by delicate deviations of the notated musical score. Therefore, expressive music performance research is aimed at establishing why, where and how these deviations take place in a piece of music. To build computational models of expressive performance is to connect the properties of a musical score and performance context with the physical parameters of a performance, such as timing, loudness, tempo, articulation and so on. These models help us to gain a better understanding of expressive music performance and provide tools to perform music by machine. Different strategies have been employed in expressive performance research (e.g., analysis-by-measurement, analysis-by-synthesis, machine learning and so on) in order to capture common performance principles or discover diversity. Comprehensive reviews about these works can be found in [12, 16]. As a matter of fact, social factors, including the influence of historical practices and the interactions between performers and audience, play an important role in music performance [6]. More precisely, our hypothesis is that expressive music performances emerge through interaction and development in a society of performers and listeners. However, the frequently used strategies can help little to investigate this aspect. Therefore, our research is aimed at a model that is able to develop performance strategies, with a

society of artificial agents. Another possible contribution of this approach is to explore the effect of social learning and evolution, which has attracted a great deal of interest in field such as artificial life, evolutionary computation and epigenetic robotics. We consider music as a natural test-field of unique research value. The work developed so far includes studying the effect of generational inheritance or social interaction. Reported earlier [19] was a Genetic Algorithm (GA)-based system that generates suitable performance profiles, through generational inheritance and mutation. This paper presents further developments, where the agents develop their performance strategies (also referred to as “performance profiles”) by imitating each other. Performance profiles are represented as hierarchical pulse sets, defining the deviations for notes’ duration and amplitude values when playing a piece of music (in MIDI format). A set of fitness rules associated with information about the structure of the piece of music in question has been devised for agent performers to evaluate the performances.

2 Music Performance with Hierarchical Pulse Set

2.1 Notion of pulse set

Fig. 1a shows a pulse represented as a curve of measurements of finger pressure on a pressure sensor pad. The information in a pulse is a wrap of specific temporal patterns with amplitude patterns, and can be quantified (width and height correspond to duration and amplitude, respectively), as depicted in Fig 1b. A pulse can operate at different levels of temporal organization and can be grouped into a hierarchical structure [5]. Clynes proposed to represent a hierarchical pulse set as a matrix of duration and amplitude values (shown in Fig 1c), which defines the deviations of the physical attributes of musical notes. We obtain performance generated by machine when a computer is used to modulate the physical attributes of musical notes following these deviations.

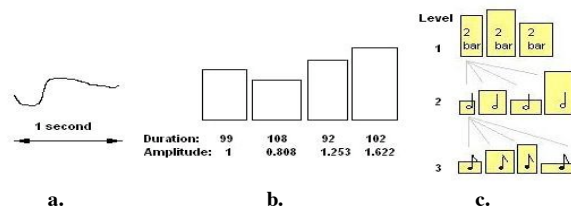


Fig. 1. Illustration of a pulse and the notion of hierarchical pulse sets. (After Clynes.)

- (a) A pulse represented as finger pressure measurements in time.
- (b) A representation of a pulse as a wrap of real numbers (duration vs. amplitude).
- (c) A hierarchical pulse set derived from grouping pulses.

2.2 Pulse sets as performance profiles

We adopted the notion of hierarchical pulse sets to represent performance profiles in our work for three reasons. Firstly, the choices of notes’ duration and amplitude significantly influence the expressive quality of a music performance. Secondly, the hierarchical nature of pulse sets matches important features of most music genres; e.g., the notions of grouping and hierarchical structures. Finally, we regard

hierarchical pulse sets as rather compact and informative forms for generative music interpretation.

2.2.1 Representation of a pulse set

Table 1 shows an example of a hierarchical pulse set and the meaning of its components. This example is the quantification of the pulse set shown in Fig. 1c.

Table 1. Representation of Pulse Set and Explanation

Pulse set example	Meaning
8	The length of note at the lowest level (8 th note)
4 4 3	Number of elements in each level (from the lowest level to the highest)
0.539 0.762 0.953 1.119	Level 3 Amplitude
73 93 106 124	Level 3 Duration
0.853 0.798 0.998 1.333	Level 2 Amplitude
92 103 114 118	Level 2 Duration
1.398 1.476 1.464	Level 1 Amplitude
109 121 90	Level 1 Duration

2.2.2 Calculating a deviation pattern from a pulse set

In our experiments, musical pieces are stored in a numerical format of our own design. An excerpt is illustrated in Fig.2, together with its respective score.



Fig. 2. Representation of a music excerpt in numerical form (left hand side) and its respective standard music notation (right hand side).

The example in Table 1 defines a performance profile containing 48 ($4 \times 4 \times 3$) elements in total. Each element's duration and amplitude deviation values are calculated in a top down manner, by multiplying parameters of related elements in different hierarchical levels. For instance, the 1st and the 40th pulse element (represented as e_1 , e_{40}) in this list are related with the following elements of the pulse set:

- e_1 : the 1st in Level 1, the 1st in Level 2, the 1st in Level 3
- e_{40} : the 3rd in Level 1, the 2nd in Level 2, the 4th in Level 3

Table 2. Calculation for a pulse element in a pulse set.

Note	Duration	Amplitude
e_1	$73 \times 92 \times 109 / 100^3$	$1.398 \times 0.853 \times 0.539$
e_{40}	$90 \times 103 \times 124 / 100^3$	$1.464 \times 0.789 \times 1.119$

According to the values shown in Table 1, we obtain deviations of e_1, e_{40} as listed in Table 2. Once started, the deviation patterns are repeated until the piece finishes.

2.2.3 *Manipulating the notes' physical parameters*

In order to calculate the physical parameters of the notes of a piece (that is, their playback duration and loudness) further steps are needed after the deviation patterns are calculated. Firstly, we look up a note's start time T_n in the aforementioned duration and loudness deviation lists. And then we apply a method similar to Clynes' method [4]: a note's playing time D_n (if this note is longer than the smallest unit) is calculated by summing up the durations of all the occupied pulse components, and its amplitude A_n is identical to the amplitude value of its first pulse component. After the modifications, the system produces a new MIDI file with expressions added on. In the following section, we introduce the evaluation of such a performance according to structural rules.

3 Fitness Function based on Musical Structure

The strong relation between music structure and expression in music performance is emphasized enormously in the literature [3, 13, 14]. This helps to explain the stability of performances by the same person over a number of years, as well as the commonality shared by different performers. On the other hand, dissimilar interpretations of the same piece may be partially explained by varied understanding of musical structure and also different approaches to signify certain structural features. As a matter of fact, the connection between musical structure and performance features is a prerequisite for modelling expressive music performance. To discover and explain performances' commonality and diversity has always been a key topic in expressive performance research.

Our approach starts by taking the musical score of the piece in question with no information about performance. However, some performance principles are needed to roughly guide the agents' interaction and evolution. Rather than a comprehensive collection of fixed performance rules, we established a few basic and flexible principles. In the present system, these principles are used by the agents to evaluate a performance, and the way to incorporate them is equivalently predefined for all agents. For the future, we plan to give freedom to each agent so that they may develop combinations of rules dynamically to evaluate a performance.

Our descriptive performance principles are explained in the following paragraphs. These principles are then used by each agent to compare its own performance with the performances of others, determining whether to imitate or not.

3.1 Selected musical structures

Performers communicate musical idea to listeners through their performances. We regard grouping and accentuation as the most important principles to facilitate this communication, very much like in human speech. In other words, we prefer performance profiles that properly highlight grouping and accentuation structure in music. In terms of tonal music, three components - rhythm, melody and harmony -

have all been taken into account. To break down the problem, we analyse the development of each aspect in the piece, and combine the cues from them that contribute to grouping and accentuation. The overall analysis contains two parts corresponding to both timing and amplitude deviations. Furthermore, the analysis is refined to note level, compatible with the deviation pattern decided by a pulse set. The method is shown in detail below.

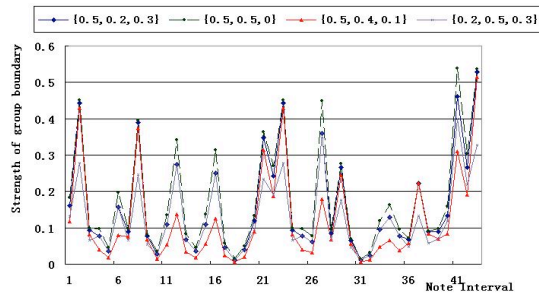


Fig. 3. Strength of group boundary calculated as a weighted sum of inter-onset intervals, pitch intervals and harmonic distances. Different lines plot the resultant strength when assigning different weight to each element.

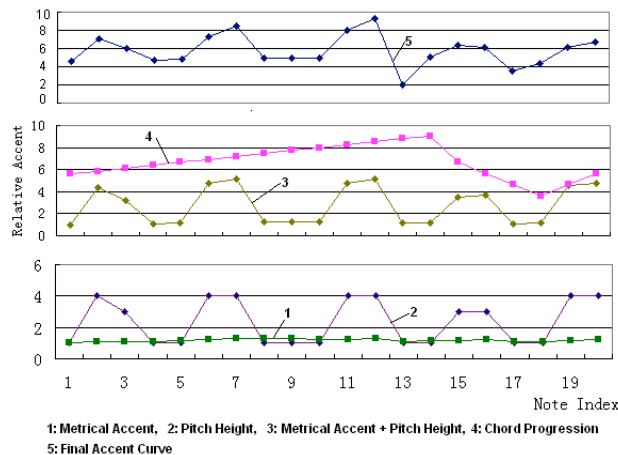


Fig. 4. Accentuation pattern. Calculated as multiplication of related elements (metrical accent, pitch height and chord progression) in corresponding positions.

3.2 Structure Analysis

In the present system, the grouping analysis uses a modified version of Cambouropoulos' Local Boundary Detection Model (LBDM) [1]. The model calculates boundary strength values for every two consecutive notes in a melodic surface. The higher value represents the higher strength of local discontinuities. In our system, we take into account the degree of change related to a note's time, pitch and harmony intervals. The latter was inspired by the "melodic charge" rule devised by

Friberg and Sundberg [14]. We assign a distance value to each note according to the probe tone ratings, given by Krumhansl [8] (Fig. 3).

In terms of accentuation pattern, we have combined metrical accentuation, melody contour and chord progressions. The positions of chord progression are calculated using Temperley's software Melism [15], and the value associated to each chord is again referred to Krumhansl [8]. The method to combine these aspects is inspired by a multi-level learning strategy proposed by Widmer and Tobudic [17]. That is, the final accentuation analysis is calculated through multiplying the corresponding value for each note in every analysis. This process is illustrated in Fig. 4.

3.3 Selected performance principles

Our descriptive performance principles are very much inspired by Clarke's generative rules for expressive performance [2], which are:

- *Principle 1*: to indicate the structural direction by parametric gradients.
- *Principle 2*: to indicate group structures by parametric continuities and discontinuities.
- *Principle 3*: to accentuate individual events by local parametric (tempo/amplitude) intensification or contrast.

Another important rule concerns the tempo curve (phrase arch) within a phrase, as discussed by a number of experts on the topic of music performance [7, 9, 10, 13].

- *Principle 4*: a phrase is always performed with an initial accelerando and subsequent ritardando.

These four principles address our needs very well, with their suitable covering of grouping and accentuation features, and being descriptive rather than quantified. Those words such as “gradients”, “(dis)continuities”, “contrast” make sense in the context of available structural analysis - curve or contour for both duration and amplitude - referred to as $Curve_{(dur)}$ and $Curve_{(amp)}$, respectively. Therefore, we implemented them with several rules that indicate values. And the total fitness of a pulse set is based on how it fulfills these rules, consisting of two components, Fit_{DurDev} and Fit_{AmpDev} respectively representing the fitness of timing (duration deviation) and dynamic (amplitude deviation).

Duration Deviation: Fit_{DurDev} is obtained with three rules. Generally speaking, since it is a common expressive option to slow down when the music is approaching its group boundary, it's preferred that the note closer to a group boundary has higher value for $Curve_{(dur)}$. However, over lengthening of notes are not encouraged. These rules are formalized as follows:

Rule 1: For every two consecutive nodes of $Curve_{(dur)}$, d_m, d_{m+1} , the related notes n_d and n_{d+1} have duration deviation $durDev_m, durDev_{m+1}$, then value v_1 is returned as follows:

$$v_1 = \sum_{m=0}^{N_{note}-2} \begin{cases} 1 & (\Delta d \cdot \Delta durDev > 0) \\ 0 & (\Delta d \cdot \Delta durDev = 0) \\ -1 & (\Delta d \cdot \Delta durDev < 0) \end{cases} \quad \text{where} \quad \begin{aligned} \Delta d &= d_{m+1} - d_m, \\ \Delta durDev &= durDev_{m+1} - durDev_m \end{aligned}$$

Rule 2: For all local peak nodes c_p on $Curve_{(dur)}$, if $durDev_p < durDev_{p-1}$ and $durDev_p < durDev_{p+1}$, then v_2 is increased by 1.

Rule 3: Given a preset maximum deviation Max_{durDev} , the pulse set should be punished if a note has larger deviation than $\pm Max_{durDev}$. That is, increase v_3 by 1.

The value of Fit_{DurDev} depends on how well a pulse set obeys the first rule and how often the violation of the last two rules takes place:

$$Fit_{DurDev} = v_1 / N_{note-1} - v_2 / N_{peak} - v_3 / N_{note}$$

where N_{note} is the total number of notes in the piece and N_{peak} is the number of local boundaries.

Amplitude Deviation: The rule to decide Fit_{AmpDev} is similar to *Rule 1*. It is an evaluation of how well the notes' amplitude fits the accentuation analysis $Curve_{(amp)}$.

Rule 4: For every two consecutive nodes of $Curve_{(amp)}$, a_m, a_{m+1} , the related notes n_m and n_{m+1} have $ampDev_d, ampDev_{d+1}$, then v_4 is calculated as follows:

$$v_4 = \sum_{m=0}^{N_{note}-2} \begin{cases} 1 & (\Delta a \cdot \Delta ampDev > 0) \\ 0 & (\Delta a \cdot \Delta ampDev = 0) \\ -1 & (\Delta a \cdot \Delta ampDev < 0) \end{cases} \quad \text{where} \quad \begin{cases} \Delta a = a_{m+1} - a_m, \\ \Delta ampDev = ampDev_{m+1} - ampDev_m \end{cases}$$

Total fitness

In the present version of our system, we define the total fitness of a pulse set to be the weighted sum of Fit_{DurDev} and Fit_{AmpDev} . That is, $Fitness = w * Fit_{DurDev} + Fit_{AmpDev}$. We set w to be larger than 1, to differentiate the listeners' higher ability to detect short timing changes than to detect loudness changes.

4 Imitation

In each generation, agent performers listen to performances by other agents, evaluate them, and learn from those having better performance than their own performances. In previous sections, we explained how an agent performer plays a piece and evaluates a performance. The following paragraphs explain the process whereby an agent performer imitates the performance of another agent. Imitation, as a form of social learning, is an effective way of taking advantage of the knowledge of others, and of particular value in circumstances where supervised training data is not applicable [18]. This is the case with our system, where we expect performance profiles emerging through agents' interaction. The imitation procedure is as follows: in one generation, each agent performer selects a few performances that it will listen to. Firstly, all performances are pre-classified, in order to reduce the chance that a performer is exposed to peers that play much worse than itself. In the present system,

agent performers share the measurement of a given performance. Performances are classified into three groups according to this measurement: top, middle or low. In each group, there are 5, 20 and 25 agents, respectively, from a population of 50 agents in total. Each performer selects 5 other agents (i.e., performances) to listen to. Table 3 shows the distribution of selected performances for individuals in different groups. Within one group, the selection is taken randomly.

Table 3. Distribution of imitation objects

	Low	Middle	Top
Low	2	2	1
Middle	0	3	2
Top	0	0	5

When an agent performer Pr0 listens to a performance by another agent Pr1, it evaluates its own performance P0 and the performance P1 by the other agent, getting the evaluation f_0 and f_1 respectively. If $f_0 > f_1$, Pr0 simply ignores P1. Otherwise, Pr0 imitates Pr1 by modifying the interpretation (in terms of timing and loudness) for the notes in P0. The details of the modification are not introduced here due to lack of space; it suffices to say that this is done firstly at group-level and then at note-level.

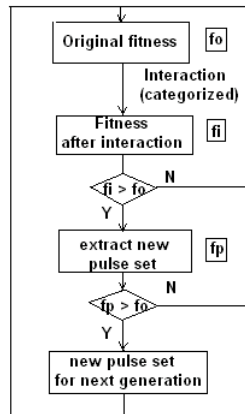


Fig. 5. Interaction procedure of an agent performer in the society

Particular attention is paid to the notes at group boundaries or accentuated beats. If an agent selects itself by chance, it performs self-mutation instead of imitation. When all imitations are completed each performer will sport a new performance profile (or pulse set). For example, taking performer Pr0, it will sport a new performance P0' and the evaluation will be f_i . If $f_i < f_0$, then Pr0 will keep its original pulse set. Otherwise, if one's performance has improved after imitation or self-mutation, an introspection process will be applied to extract a new hierarchical pulse set from this new performance. However, note that the coding of a performance profile into a hierarchical pulse set is not a trivial task with promised fidelity. We search for the matching pulse set of a given performance profile by means of a genetic algorithm. And this new pulse set replaces the original if appropriate. The procedure is depicted in Fig 5.

5 Demonstration

As a demonstration, we present an example with a melody from Robert Schumann's *Träumerei*. Fig.6 shows the structural analysis of the melody used for calculating the fitness, including metrical analysis (the numbers at the bottom of the notes) and harmonic progression (with the chord names above the staves). We conducted a number of experiments, each of which involving 50 individual pulse sets randomly generated for the first generation. Each experiment lasted for 100 generations (imitations). From each run we have recorded the following data: (a) Interaction result: for every agent performer, information about which group (low, middle, top) it belongs to, with whom it interacts and their categories, and the outcome of the interaction (that is, whether the fitness increases or decreases); (b) Fitness value: in a generation, for every agent performer, its original fitness, fitness after imitation, and the fitness value of the extracted pulse set; (c) Modified performance profiles after interaction, and the performance generated by the new extracted pulse set; (d) Extracted pulse sets as individuals for the next generation.

The figure shows three staves of musical notation for the soprano part of Robert Schumann's *Träumerei*. The first staff begins at measure 1 and ends at measure 11. The second staff begins at measure 6 and ends at measure 11. The third staff begins at measure 11 and ends at measure 11. The melody is written in a treble clef with a key signature of one flat (B-flat). The time signature is 4/4. The score includes various chords and metrical patterns indicated by numbers below the notes. The chords are labeled as I, IV⁶, V, V₆, and i₄⁶. The metrical patterns are represented by numbers 1, 2, 3, 4, and combinations thereof, indicating the pulse sets used for fitness calculation.

Fig. 6. An excerpt of the soprano part from Robert Schumann's *Träumerei*.

One of our main goals in examining the results is to ascertain whether the interactions between the agent performers (in the form of imitation) improve the overall fitness of the society. This will be analyzed in the following paragraphs. Also, we have compared the result obtained when: (1) agents only interact with one another, and (2) agents also inherit information, in addition to interact with one another. However, the latter will not be discussed in this paper due to lack of space.

5.1 Comparison between original fitness and fitness after interaction

Fig.7 shows the minimum, average and maximum fitness values before and after the interactions, for a run of 100 generations. From the figure, we can easily observe that interaction increases the overall fitness of the agent society as a whole. This is more obvious comparing the minimum and average fitness of all the agents, before and after the interactions. It is fair to say that interaction among individuals diminishes differences within the society, by effectively improving the worse performances.

However, it is also clear that the society's best performance doesn't improve significantly through interaction. This is not surprising: firstly, the guidance that a top performer can possibly get during either imitation or self-mutation is comparably

weaker than the guidance that poor performers can get. Secondly, the biases of extracting pulse set from modified performance profiles sometimes counteract an agent's achievement through interaction. This is a factor that oscillates the best performances through generations. And it is of interest here to examine the effect of "introspection" in more detail; we are currently working on this.

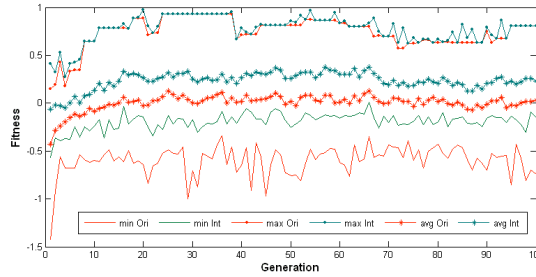


Fig.7. Original fitness and fitness after interaction in each generation

5.2 Interaction between agents

At each generation, we have recorded the percentage of interactions through which the imitator has increased its fitness. Other information includes which group (low, mid, high) the imitator belongs to and the group of those being imitated. As shown in Fig. 8, to imitate those belonging to a better group than its own group is more likely to improve an imitator's own fitness. Again, this is not surprising; the larger the ranking difference, the more likely it is to result in an improvement.

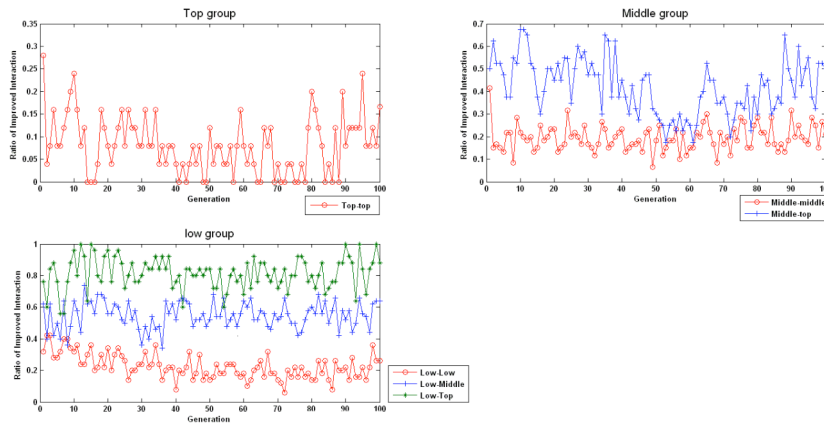


Fig. 8. Interaction result. The ratio of interactions that improve an agent's fitness

5.3 Diversity among evolved agent performers

Fig.9 is an example of the fitness distribution at the end of a run, with the pulse set of the circled agents shown in Table 4. Here we take Agent 9, Agent 16 and Agent 26 for example, because their fitness values are close to each other. Fig.10 draws the performance profiles determined by each of these 3 agent performers. They are quite different from each other, especially with respect to tempo deviation. However, similarity is found with local segments of notes.

We have previously reported a model for evolving suitable pulse set using GA for interpreting a piece of music [19]. The diversity and commonality among evolved performance profiles were observed. Compared with the previous GA system, the models shown in this paper is much slower to converge, either in terms of performance profiles or pulse sets.

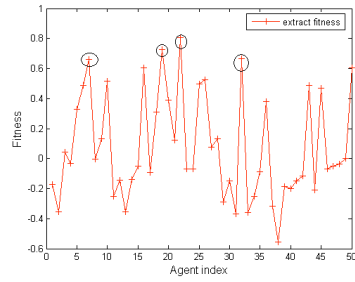


Fig. 9. Fitness of all agents in Generation 99
Agent 7,19,22,32 are selected for comparison

Agent 7	Agent 19
8	4
5 5 5 2	6 4 4 2
0.772 0.907 0.956 1.127 1.158	0.958 0.623 0.659 0.887 1.408 1.303
79 89 123 124 130	131 74 100 63 143 92
0.913 1.327 0.817 0.535 1.128	1.328 1.298 0.581 0.674
65 113 54 105 83	133 123 137 132
1.317 1.224 1.11 1.026 0.523	1.397 0.728 0.815 1.458
122 103 59 75 141	138 100 145 69
0.521 1.263	0.735 1.457
142 75	91 135
Agent 22	Agent 32
16	8
5 5 6 3	5 5 5 2
1.342 0.645 0.88 1.232 0.5	1.419 1.275 0.917 0.649 0.895
139 144 56 68 149	63 68 91 76 60
0.806 1.383 0.941 0.925 0.669	1.163 0.556 1.212 1.432 1.064
95 146 138 128 69	143 119 76 86 112
1.081 1.225 0.565 1.447 0.58 1.272	0.729 0.806 1.321 1.013 1.335
149 87 80 52 63 129	70 126 145 118 108
0.859 0.626 0.638	1.092 0.596
105 88 77	130 133

Table 4. Pulse set of Agent 7, 19, 22, 32

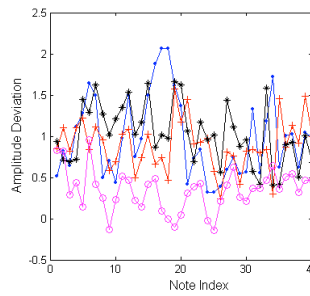
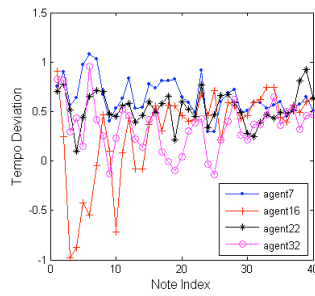


Fig. 10. Performance profiles (Tempo/Amplitude deviation) determined by different pulse set having close fitness

6 Concluding Remarks and Ongoing Work

In this paper we introduced an artificial agent society that evolves pulse sets (or performance profiles) for musical performance through interaction between agents. Fitness rules derived from the structure of the piece to be performed are used to evaluate the evolved performances. The evolved performances were examined both objectively (by comparing the figures of deviation patterns by different pulse sets) and subjectively (by listening to the “interpreted” MIDI files). As explained earlier, we believe that the dynamic characteristic of the system is largely determined by the way in which we combine aspects of structural analysis forming a fitness function. We are currently studying the possibility of specifying different fitness weights for each agent performer. Also, we are interested to establish whether common and suitable combinations of performance rules can emerge through the interactions.

References

- [1] Cambouropoulos E. "The Local Boundary Detection Model (LBDM) and its Application in the Study of Expressive Timing". In Proceedings of the International Computer Music Conference (ICMC'2001) 17-22 September, Havana, Cuba.
- [2] Clarke, E. F. "Generative Processes in Music". The Psychology of Performance, Improvisation, and Composition. Oxford Science Publications (1988).
- [3] Clarke, E. F. "Expression and communication in musical performance". (1988).
- [4] Clynes, M. "Generative principles of musical thought integration of microstructure with structure". Journal For The Integrated Study Of Artificial Intelligence, Cognitive Science And Applied Epistemology 3 (1986) 185-223.
- [5] Clynes, M. "Microstructural musical linguistics: composers' pulses are liked most by the best musicians". Cognition. International Journal of Cognitive Science, 55, (1995), 269-310.
- [6] Davidson, J. W. and North, A. C. The Social Psychology of Music. Oxford University Press (2006).
- [7] Friberg A. "Matching the rule parameters of Phrase Arch to performances of 'Traumerei': a preliminary study". Proceedings of the KTH symposium of Grammars for Music Performance, May 1995
- [8] Carol Krumhansl. Cognitive Foundations of Musical Pitch. Oxford University Press, Oxford, 1991.
- [9] Neil Todd, "A computational model of rubato". Contemporary Music Review, 3, (1989), 69-88.
- [10] Neil Todd, "The dynamics of dynamics: A model of musical expression". Journal of the Acoustical Society of America, 91(6) (1992).
- [11] Parncutt, R 1997. "Accents and Expression in Piano Performance," in W. Auhagen et al (Eds), Systemische Musikwissenschaft. Krefeld: Dohr.
- [12] Poli, G. D. "Methodologies for expressiveness modelling of and for music performance". Journal Of New Music Research 33 (2004) 189-202.
- [13] Repp, B. H. "Diversity and commonality in music performance: An analysis of timing microstructure in Schumann's Träumerei". Journal of the Acoustical Society of America (92).
- [14] Sundberg, J. "Grouping and Differentiation Two Main Principles of Music", Integrated Human Brain Science L Theory, Method Application (Music) T. Nakad (Ed). (2000)
- [15] Temperly, D. The Cognition of Basic Musical Structures. The MIT Press (2004).
- [16] Widmer, G. and Goebel, W. "Computational models of expressive music performance: The state of the art". Journal of New Music Research 33 (2004) 203-216.
- [17] Widmer G., Tobudic A. "Playing Mozart by Analogy: Learning Multi-level Timing and Dynamics Strategies" Journal of New Music Research, Volume 32, Number 3, September 2003, pp. 259-268(10)
- [18] Kerstin Dautenhahn, Christopher L. Nehaniv, and Aris Alissandrakis. "Learning by Experience from Others — Social Learning and Imitation in Animals and Robots" In R. Kuhn, et al, Eds., "Perspectives on Adaptivity and Learning: An Interdisciplinary Debate", pp. 217-241, Springer Verlag.
- [19] Zhang, Q. and Miranda, E. R. "Evolving Musical Performance Profiles using Genetic Algorithms with Structural Fitness", Proceedings of Genetic and Evolutionary Computation Conference (GECCO 2006), Seattle (USA).