Towards Artificial Creativity: Examples of some applications of AI to music performance

Ramón LOPEZ DE MANTARAS
IIIA, Artificial Intelligence Research Institute
CSIC, Spanish Scientific Research Council
Campus UAB 08193 Bellaterra, Spain

Abstract. Creativity is an important element of our problem solving capabilities that involves, among others, heuristic search, reasoning, analogy, learning, and reasoning under constraints. In this paper I describe examples of computer programs capable of replicating some aspects of creative behavior in the domain of music

Keywords. Artificial Intelligence, Creativity, Music

Introduction

New technologies, and in particular Artificial Intelligence, are drastically changing the creative processes. Computers are playing very significant roles in creative activities such as music, architecture, fine arts and science. Indeed, the computer is already a canvas, a brush, a musical instrument, etc. However, I believe that we must aim at more ambitious relations between computers and creativity. I mean relations other than just seeing the computer as a tool to help human creators but a creative entity in itself.

In this paper I will, therefore, briefly address the question of the possibility of building creative machines. But, what is creativity? Why is it so mysterious?. Creativity seems mysterious because when we have creative ideas it is very difficult to explain how we got them and the best we can do is talk about "inspiration", "intuition" when we try to explain creativity. The fact that we are not conscious about how a creative idea manifests itself does not necessarily imply that a scientific explanation cannot exist. As a matter of fact, we are neither aware of how we perform other activities such as language understanding, pattern recognition, etc. but we have better and better AI techniques able to replicate such activities.

For many, creativity requires the possession of a mysterious and unusual "gift" that cannot be explained. This is possibly due to the fact that, often, creativity is associated with geniuses like, Bach or Picasso, but should the term "creative" be only applied to a few men and women? Or is it rather one more aspect of intelligence? In other words, is creativity the result of a very special mechanism only existing in very unusual minds or is rather one more aspect of our problem solving activity? An operational, and widely accepted, definition of creativity is "a creative idea is a novel and valuable combination of known ideas". In other words, new physical laws, new theorems, new musical pieces can be generated from a finite set of existing elements. I agree with this definition and, therefore, with those that believe that creativity is an advanced form of our problem solving capabilities that involves, among others, memory, heuristic search, reasoning by analogy, learning and reasoning under constraints and therefore it should be possible to replicate by means of computers.

Based on this belief, in this paper I give some examples of representative applications to expressive music performance that employ AI techniques.

For further reading I recommend the books by Boden [1,2], Dartnall [3], Partridge & Rowe [4], and Bentley & Corne [5]; as well as the papers by Rowe & Partridge [6], and Buchanan [7].

Synthesizing expressive music

One of the main limitations of computer-generated music has been its lack of expressiveness, that is, lack of "gesture". Gesture is what musicians call the nuances of performance that are unique and subtly interpretive or, in other words, creative.

One of the first attempts to address expressiveness in music is that of Johnson [8]. She developed an expert system to determine the tempo and the articulation to be applied when playing Bach's fugues from "The Well-Tempered Clavier". The rules were obtained from two expert human performers. The output gives the base tempo value and a list of performance instructions on notes duration and articulation that should be followed by a human player. The results very much coincide with the instructions given in well known commented editions of "The Well-Tempered Clavier". The main limitation of this system is its lack of generality because it only works well for fugues written on a 4/4 meter. For different meters, the rules should be different. Another obvious consequence of this lack of generality is that the rules are only applicable to Bach fugues.

The work of the KTH group from Stockholm [9, 10, 11, 12] is one of the best known long term efforts on performance systems. Their current "Director Musices" system incorporates rules for tempo, dynamic, and articulation transformations constrained to MIDI. These rules are inferred both from theoretical musical knowledge and experimentally by training, specially using the so-called analysis-by-synthesis approach. The rules are divided in three main classes: *Differentiation rules*, which enhance the differences between scale tones; *Grouping rules*, which show what tones belong together; and *Ensemble rules*, that synchronize the various voices in an ensemble.

Canazza et al [13] developed a system to analyze how the musician's expressive intentions are reflected in the performance. The analysis reveals two different expressive dimensions: one related to the energy (dynamics) and the other one related to the kinetics (rubato) of the piece. The authors also developed a program for generating expressive performances according to these two dimensions.

The work of Dannenberg and Derenyi [14] is also a good example of articulation transformations using manually constructed rules. They developed a trumpet synthesizer that combines a physical model with a performance model. The goal of the performance model is to generate control information for the physical model by means of a collection of rules manually extracted from the analysis of a collection of controlled recordings of human performance.

Another approach taken for performing tempo and dynamics transformation is the use of neural network techniques. In [15] a system that combines symbolic decision rules with neural networks is implemented for simulating the style of real piano performers. The outputs of the neural networks express time and loudness deviations. These neural networks extend the standard feed-forward network trained with the back propagation algorithm with feedback connections from the output neurons to the input neurons.

We can see that, except for the work of the KTH group that considers three expressive resources, the other systems are limited to two resources such as rubato and dynamics, or rubato and articulation. This limitation has to do with the use of rules. Indeed, the main problem with the rule-based approaches is that it is very difficult to find rules general enough to capture the variety present in different performances of the same piece by the same musician and even the variety within a single performance [16]. Furthermore, the different expressive resources interact with each other. That is, the rules for dynamics alone change when rubato is also taken into account. Obviously, due to this interdependency, the more expressive resources one tries to model, the more difficult is finding the appropriate rules.

We have developed a system called SAXEX that won the Best Paper Award at the 1997 International Computer Music Conference [17]. SAXEX is a computer program capable of synthesizing high quality expressive tenor sax solo performances of jazz ballads based on cases representing human solo performances. Previous rule-based approaches to that problem could not deal with more than two expressive parameters (such as dynamics and rubato) because it is too difficult to find rules general enough to capture the variety present in expressive performances. Besides, the different expressive parameters interact with each other making it even more difficult to find appropriate rules taking into account these interactions.

With CBR, we have shown that it is possible to deal with the five most important expressive parameters: dynamics, rubato, vibrato, articulation, and attack of the notes. To do so, SaxEx uses a case memory containing examples of human performances, analyzed by means of spectral modeling techniques and background musical knowledge. The score of the piece to be performed is also provided to the system. The heart of the method is to analyze each input note determining (by means of the background musical knowledge) its role in the musical phrase it belongs to, identify and retrieve (from the case-base of human performances) notes with similar roles, and finally, transform the input note so that its expressive properties (dynamics, rubato, vibrato, articulation, and attack) match those of the most similar retrieved note. Each note in the case base is annotated with its role in the musical phrase it belong to as well as with its expressive values. Furthermore, cases do not contain just information on each single note but they include contextual knowledge at the phrase level. Therefore, cases in this system have a complex object-centered representation.

Although limited to monophonic performances, the results are very convincing and demonstrate that CBR is a very powerful methodology to directly use the knowledge of a human performer that is implicit in her playing examples rather than trying to make this knowledge explicit by means of rules. Some audio results can be listen at www.iiia.csic.es/arcos/noos/Demos/Aff-Example.html. More recent papers by Arcos and Lopez de Mantaras [18] and by Lopez de Mantaras and Arcos [19], describe this system in great detail.

Based on the work on SaxEx, we have decelopped TempoExpress [20] a case-based reasoning system for applying musically acceptable tempo transformations to monophonic audio recordings of musical performances. TempoExpress has a rich description of the musical expressivity of the performances, that includes not only timing deviations of performed score notes, but also represents more rigorous kinds of expressivity such as note ornamentation, consolidation, and fragmentation. Within the tempo transformation process, the expressivity of the performance is adjusted in such a way that the result sounds natural for the new tempo. A case base of previously performed melodies is used to infer the appropriate expressivity. The problem of changing the tempo of a musical performance is not as trivial as it may seem because it involves a lot of musical knowledge and creative thinking. Indeed, when a musician performs a musical piece at different tempos the performances are not just time-scaled versions of each other (as if the same performance were played back at diferent speeds). Together with the changes of tempo, variations in musical expression are made [21]. Such variations do not only affect the timing of the notes, but can also involve for example the addition or deletion of ornamentations, or the consolidation/fragmentation of notes. Apart from the tempo, other domain specific factors seem to play an important role in the way a melody is performed, such as meter, and phrase structure. Tempo transformation is one of the audio post-processing tasks manually done in audio-labs. Automatizing this process may, therefore, be of industrial interest.

Other applications of CBR to expressive music are those of Suzuki et al. [22], and those of Tobudic and Widmer [23, 24]. Suzuki et al. [22], also use examples cases of expressive performances to generate multiple performances of a given piece with varying musical expression, however they deal only with two expressive parameters. Tobudic and Widmer [23] apply instance-based learning (IBL) also to the problem of generating expressive performances. The IBL approach is used to complement a note-level rule-based model with some predictive capability at the higher level of musical phrasing. More concretely, the IBL component recognizes performance patterns, of a concert pianist, at the phrase level and learns how to apply them to new pieces by analogy. The approach produced some interesting results but, as the authors recognize, was not very convincing due to the limitation of using an attribute-value representation for the phrases. Such simple representation does not allow taking into account relevant structural information of the piece, both at the sub-phrase level and at the inter-phrasal level. In a subsequent paper, Tobudic and Widmer [24], succeeded in partly overcoming this limitations by using a relational phrase representation.

The possibility for a computer to play expressively is a fundamental component of the so-called "hyper-instruments". These are instruments designed to augment an instrument sound with such idiosyncratic nuances as to give it human expressiveness and a rich, live sound. To make an hyper-instrument, take a traditional instrument, like for example a cello, and connect it to a computer through electronic sensors in the neck and in the bow, equip also with sensors the hand that holds the bow and program the computer with a system similar to SAXEX that allows to analyze the way the human interprets the piece, based on the score, on musical knowledge and on the readings of the sensors. The results of such analysis allows the hyper-instrument to play an active role altering aspects such as timbre, tone, rhythm and phrasing as well as generating an accompanying voice. In other words, you have got an instrument that can be its own intelligent accompanist. Tod Machover, from MIT's Media Lab, constructed such an hypercello and the great cello player Yo-Yo Ma premiered, playing the hypercello, a piece, composed by Tod Machover, called "Begin Again Again..." at the Tanglewood Festival several years ago.

Apparently or really creative?

The main limitation of computational models of creativity is the absence of built-in evaluation criteria to select the most valuable new combinations among the numerous generated ones. This is especially difficult in artistic domains such as music. In these domains the selection is generally done by humans. This is indeed the case of our system SaxEx and of the other systems we know, particularly those based on evolutionary computation. In other domains there have been some attempts to provide the system with evaluation capabilities.

Margaret Boden pointed out that even if an artificially intelligent computer would be as creative as Bach or Einstein, for many it would be just apparently creative but not really creative. I fully agree with Margaret Boden in the two main reasons for such rejection. These reasons are: the lack of intentionality and our reluctance to give a place in our society to artificially intelligent agents. The lack of intentionality is a direct consequence of Searle's Chinese room argument, which states that computer programs can only perform syntactic manipulation of symbols but are unable to give them any semantics. This critic is based on an erroneous concept of what a computer program is. Indeed, a computer program does not only manipulate symbols but also triggers a chain of cause-effect relations inside the computer hardware and this fact is relevant for intentionality since it is generally admitted that intentionality can be explained in terms of causal relations. However, it is also true that existing computer programs lack too many relevant causal connections to exhibit intentionality but perhaps future, possibly anthropomorphic, "embodied" artificial intelligences, that is agents equipped not only with sophisticated software but also with different types of advanced sensors allowing to interact with the environment, may have enough causal connections to have intentionality.

Regarding social rejection, the reasons why we are so reluctant to accept that non human agents can be creative is that they do not have a natural place in our society of human beings and a decision to accept them would have important social implications. It is therefore much simpler to say that they appear to be intelligent, creative, etc. instead of saying that they are. In a word, it is a moral but not a scientific issue. A third reason for denying creativity to computer programs is that they are not conscious of their accomplishments. However I agree with many AI scientists in thinking that the lack of consciousness is not a fundamental reason to deny the potential for creativity or even the potential for intelligence. After all, computers would not be the first example of unconscious creators, evolution is the first example as Stephen Jay Gould [25] brilliantly points out: "If creation demands a visionary creator, then how does blind evolution manage to build such splendid new things as ourselves?"

References

[1] M. Boden, *The Creative Mind: Myths and Mechanisms*, New York; Basic Books 1991.
[2] M. Boden (Ed.); *Dimensions of Creativity*, MIT Press 1994.
[3] T. Dartnall (Ed.); *Artificial Intelligence and Creativity*, Kluwer Academic Pub. 1994.
[4] D. Partridge, and J. Rowe; *Computers and Creativity*, Intellect Books, 1994.
[5] P.J. Bentley and D.W: Corne (eds.), *Creative Evolutionary Systems*, Morgan Kaufmann 2002.
[6] J. Rowe, D. Partridge, Creativity: A survey of AI approaches, *Artificial Intelligence Review* 7 (1993), 43-70.
[7] B.G. Buchanan, Creativity at the Metalevel: AAAI-2000 Presidential Address, *AI Magazine* 22:3 (2001), 13-28.
[8] M.L. Johnson. *An expert system for the articulation of Bach fugue melodies*. In *Readings in Computer Generated Music*, ed. D.L. Baggi, 41-51. Los Alamitos, Calif.: IEEE Press. 1992.

[9] R. Bresin. *Articulation rules for automatic music performance*. In Proceedings of the 2001 International Computer Music Conference 2001. San Francisco, Calif.: International Computer Music Association. 2001.

[10] A. Friberg. *A quantitative rule system for musical performance*. PhD dissertation, KTH, Stockholm. 1995.

[11] A. Friberg, R. Bresin, L. Fryden, and J. Sunberg. Musical punctuation on the microlevel: automatic identification and performance of small melodic units. *Journal of New Music Research* 27:3 (1998), 271-292.

[12] A. Friberg, J. Sunberg, and L. Fryden. Music From Motion: Sound Level Envelopes of Tones Expressing Human Locomotion. *Journal on New Music Research*, 29:3 (2000), 199-210.

[13] S. Canazza, G. De Poli, A. Roda, and A. Vidolin. *Analysis and synthesis of expressive intention in a clarinet performance*. In Proceedings of the 1997 International Computer Music Conference, 113-120. San Francisco, Calif.: International Computer Music Association. 1997.

[14] R.B. Dannenberg, and I. Derenyi. Combining instrument and performance models for high quality music synthesis, *Journal of New Music Research* 27:3 (1998), 211-238.

[15] R. Bresin. Artificial neural networks based models for automatic performance of musical scores, *Journal of New Music Research* 27:3 (1998), 239-270.

[16] R.A. Kendall, and E.C. Carterette. The communication of musical expression. *Music Perception* 8:2 (1990), 129.

[17] J.L. Arcos, R. Lopez de Mantaras, and X. Serra; Saxex: A Case-Based Reasoning System for Generating Expressive Musical Performances. *Journal of New Music Research* 27:3 (1998), 194-210.

[18] J.L. Arcos, and R. Lopez de Mantaras; An Interactive Case-Based Reasoning Approach for Generating Expressive Music. *Applied Intelligence* 14:1 (2001), 115-129.

[19] R. Lopez de Mantaras, and J.L. Arcos; AI and Music: From Composition to Expressive Performance. *AI Magazine* 23:3 (2002), 43-57.

[20] M. Grachten, J.L. Arcos, and R. Lopez de Mantaras; *TempoExpress, a CBR Approach to Musical Tempo Transformations*. In Proceedings of the 7th European Conference on Case-Based Reasoning (Eds. P. Funk and P. A. Gonzalez Calero), Lecture Notes in Artificial Intelligence 3155 (2004), 601-615.

[21] P. Desain and H. Honing. *Tempo curves considered harmful*. In "Time in contemporary musical thought" J. D. Kramer (ed.), Contemporary Music Review. 7:2, 1993.

[22] T. Suzuki, T. Tokunaga, and H. Tanaka; *A Case-Based Approach to the Generation of Musical Expression*. In Proceedings of the 16th International Joint Conference on Artificial Intelligence, Morgan Kaufmann (1999), 642-648.

[23] A. Tobudic, and G. Widmer; *Playing Mozart Phrase by Phrase*. In Proceedings of the 5th International Conference on Case-Based Reasoning (Eds. K.D. Ashley and D.G. Bridge), Lecture Notes in Artificial Intelligence 3155 (2003), 552-566.

[24] A. Tobudic, and G. Widmer; *Case-Based Relational Learning of Expressive Phrasing in Classical Music*. In Proceedings of the 7th European Conference on Case-Based Reasoning (Eds. P. Funk and P. A. Gonzalez Calero), Lecture Notes in Artificial Intelligence 3155 (2004), 419-433.

[25] S. J. Gould; Creating the Creators, *Discover Magazine*, October (1996), 42-54.