Development of Serious Games for Music Education

Adriano Baratè, Mattia G. Bergomi, Luca A. Ludovico

Laboratorio di Informatica Musicale (LIM) -Dipartimento di Informatica - Università degli Studi di Milano (Italy)
{mattia.bergomi,adriano.barate,luca.ludovico}@unimi.it

Keywords: serious games, music, education

Serious games have proved to be an effective educational tool in many fields. The first goal of this paper is to illustrate some possible applications to music and their advantages. Moreover, music can be characterized by heterogeneous multimedia contents. Among the different facets music information is made of it is worth citing music symbols, their graphical representations as scores, their audio renderings as tracks, etc. The international standard known as IEEE 1599 is an XML multilayer format for heterogeneous music contents, and describes such different aspects in an integrated and synchronized context. Making relationships among music contents explicit provides a potentially rich educational environment. Consequently, this paper discusses the concept of multilayer learning object, introduces the IEEE 1599 standard, and finally shows some applications and case studies.

for citations:
Baratè A., Bergomi M.G., Ludovico L.A.(2013), *Development of Serious Games for Music Education,* Journal of e-Learning and Knowledge Society, v.9, n.2, 89-104. ISSN: 1826-6223, e-ISSN:1971-8829

1 Introduction

In this paper we analyze the relationship between (e)-learning and playing. This first sentence is a pun: we are going to consider the relationships between learning music theory and playing a game to learn it; at the same time in this introduction, we want to briefly discuss the links among music theory study and the practical import on an instrument. In fact this is the reason that leads someone to study music theory.

A remarkable subject in Pedagogy is the fact that playing without a goal allows our mind to build new structures and new relationships between causes and effects. Konrad Lorenz described an interesting experiment conduced by Kohner (Lorenz & Kickert, 1987). This experiment involved a chimpanzee whose name was Sultan. He was asked to reach some food that was too far from him. To reach it, the monkey had to fit two sticks to build a longer one. Sultan could not do it at the beginning.

After a few minutes trial, he started playing freely with those sticks and it found that they could be matched together. Seeing this new tool he understood he could reach the food, and so he did. Now we are not trying to say that our users are kind of monkeys, which is obvious, however this good example proves that free playing allows learning without the huge weight given by a predetermined aim.

Clearly, serious games (Abt, 1970) are designed for purposes that go beyond mere amusement. The effectiveness of games as teaching tools is due to their potential to engage players. As we said before this is the moment in which the deep learning process starts. The locution "serious game" was actually used long before the introduction of computers and electronic devices into entertainment, but our concern in this article is with the application of information technology to educational processes. Empirically one can see that this experience of effective learning has a notable consequence on the practical side of playing. One of the most interesting features that any electronic device gives is its own multilayers, synesthetic nature.

The article (Zyda, 2005) provides an up-to-date definition for serious games, depicting them as mental contests, played with a computer in accordance with specific rules that uses entertainment to further government or corporate training, education, health, public policy, and strategic communication objectives.

This paper addresses a specific knowledge field, namely music education. Serious games can be applied to this domain for a number of different purposes. For instance, as shown in the following, it is possible to teach the key concepts of music theory and instrumental practice through *ad-hoc* hardware and software frameworks.

Section 2 explains the cognitive and educational benefits of serious games for music through an example. However, in order to provide a comprehensive representation of music information, a multilayer approach is desirable. Section 3 defines and investigates some key concepts, such as the ones of *layer* and *spine*, which are implemented in the IEEE 1599 standard. Following this approach, Section 4 introduces multilayer learning objects for serious games, and finally Section 5 describes a practical application of music serious games.

2 Game Tasks: Cognitive and Educational Benefits

A serious game usually has the look and feel of a game, but it is actually a simulation of real-world events or processes. Its main goal is not mere entertainment, rather a form of education through entertainment called *edutainment* (Addis, 2005).

It is well known that children develop cognitive and intellectual constructs through manipulation, so it is obvious that music has to be thought following the natural evolution of human cognitive abilities (Ramstetter *et al.*, 2010). A serious game allows treating abstract concepts as tangible entities.

For the sake of clarity, let us introduce a practical example in the music field. In modern notation, usually three clefs (namely three graphical symbols) are used to notate music, i.e. F, C, and G. Each type of clef assigns a different reference note to the line on which it is placed. In order to facilitate writing for different tessituras, any of the clefs may theoretically be placed on any of the lines of the staff. Only nine possible distinct clefs are allowed and have been historically adopted (Kurtzman, 1994): the G-clef on the two bottom lines, the F-clef on the three top lines, and the C-clef on any line of the stave except the topmost. *iClef* is a serious game which addresses the problem of correctly reading note pitches when the clef is not a common one, typically either a treble or a bass clef. The interface shows notes with no rhythmic indication, written on a staff that carries randomly selected clefs, at an increasing metronome rate. In order to help the user, smaller noteheads suggest the next notes, written in the current clef. The user has to choose the right corresponding key on a piano-like keyboard as quickly as possible. The interface is shown in Figure 1.

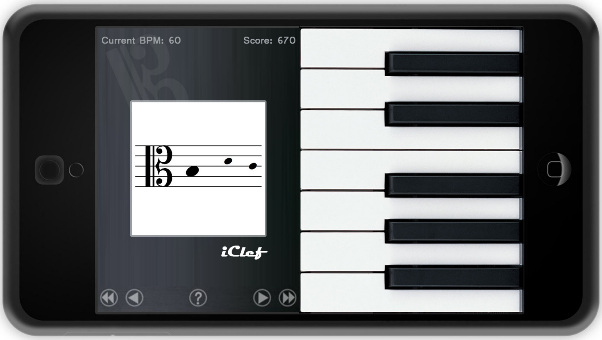
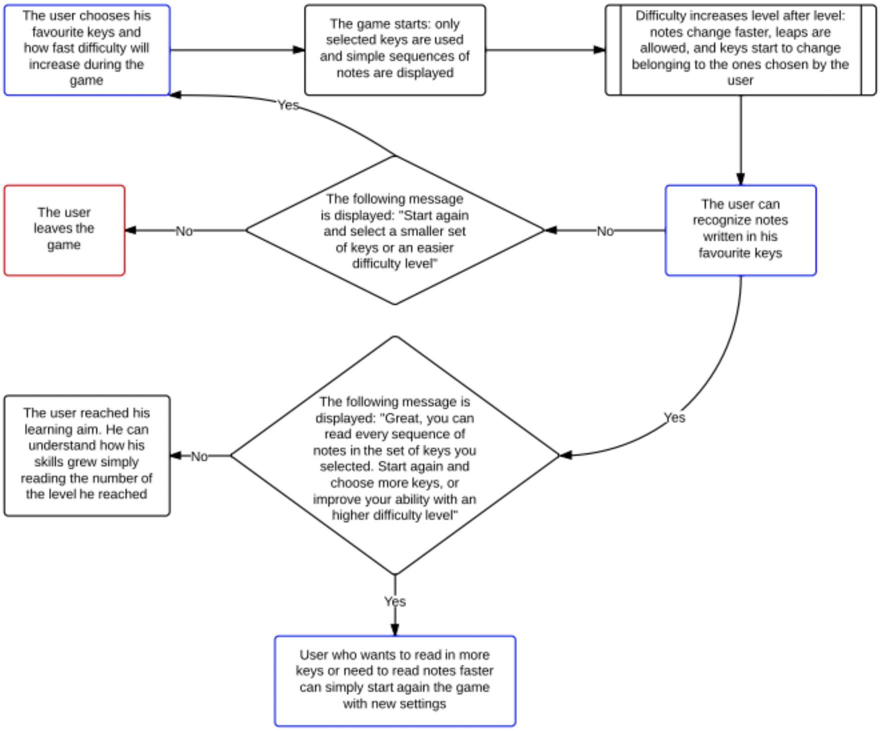


Fig. 1 - The interface of iClef.

Common elements in serious games include back story (plot/story line), game mechanics (physical functions/actions), rules (constraints), immersive environment (including 2D/3D animations), interactivity (impact of player's actions), and challenge/competition (against the game or against other players) (Antonova & Stefanov, 2011). We made a selection of those features to build a product that fits with the needs of a music student:

- Immersive environment - Notes are movable objects on a space-time graph (the sheet music) and the key changes make the game charming avoiding the boring, methodical exercise students learnt in their classroom;
- Interactivity - The user recognizes notes by touching the screen;
- Usability - Some reassurance moments are included in the game to fix notions learnt while playing and guarantee a relaxing experience of the game;
- Competition - Levels are structured with a monotonically increasing difficulty.

The fundamental task of this serious game is to make the users able to recognize notes in any key, that is one of the most boring aim a music student has to reach in any base level music course. Here follows a diagram which analyze processes and intermediate task users will reach until they reach their learning aim.



Any key represents an intermediate task, the user can evaluate himself constantly focusing his attention on the level he is playing and on the difficulty level he chose. However the user is driven to forgive his scope to simply play the game. Let us have a look to the path one will follow approaching the game.

1. The possibility to set a certain number of keys and a difficulty level is an unavoidable first step. One has to start with a low as possible stress level and have to be aware of his preexisting skills to play without a scope.
2. The following step is a really quick way to have a well-founded self-evaluation. A low difficulty level will assure a relaxing experience and a natural strengthening of preexisting and new knowledge.
3. Low stress level does not mean no stress level: at any level difficulty increases gradually. The user learns new strategies to recognize complex intervals of notes in a certain key.
4. Without loss of generality, suppose that the user chose only the key of G. In the previous steps he learnt some strategies to recognize notes. Any interval has a recognizable shape on the sheet:

A key variation corresponds to a translation of these shapes:

Now this second task of the game should be clear: the user learns some strategies working with his favorite set of keys, then he will apply them naturally to new keys. This approach assures us a high usability level and a reasonably fast pace in learning.

3 A Multilayer Approach

Almost any field of knowledge can be described from different perspectives, thus unveiling heterogeneous facets of the same entity. Undoubtedly this is the case of music. The idea of music description can embrace a number of different meanings. When people address a music work, they usually identify it through a number of metadata, e.g. its title, authors, performers, instrumental ensemble, etc. Some subsets can be very effective even if they are not complete: for instance "La Primavera by Vivaldi" is often used as a compact way to refer to "the first movement from *Le quattro stagioni*, a set of four violin concertos from Antonio Vivaldi's *Il cimento dell'armonia e dell'inventione*, Op. 8".

Moreover, a set of metadata is only one of the many ways to describe a music piece. Going back to Vivaldi example, whistling its *incipit* - even if limited to the leading voice - can be as much effective as citing its title. In fact, audio contents are another kind of description, addressing the sense of hearing. Of course, the same music piece can be performed in many different ways, as regards interpretation, score version, ensemble, etc. All these performances, even if they belong to the same category, may add further information to the overall description.

A similar approach regards the visual aspects of music. Musical notation is any system that represents scores through the use of written symbols. Many different ways are allowed for notated music, ranging from modern staff notation (Read, 1979) to neumes (Parrish, 1978), from Asian solmization (Kimiko & Yoshihiko, 1983) to Indian *sargam* (Shirali, 1977), from lute tablature (Rubsamen, 1968) to Braille (Union & Krolick, 1996). And once again, the same music piece - namely the same set of music symbols - can present different page layouts. In this case, the sense of sight is involved.

The previous ones are only few examples. Since music communication is very rich, also the number of heterogeneous descriptions is potentially high. The problem of catching and describing heterogeneity in music will be addressed in the following sections. As we will explain, a multilayer approach is fit for treating complex information entities by keeping contents properly organized within a unique framework.

3.1 The Concept of Layer

As mentioned before, our approach towards the design of music-oriented serious games starts from a *multilayer* description of music information.

In this sense, the first key concept to deepen is the idea of layer itself. Let us introduce some examples coming from different contexts. A dictionary would define a layer as a covering surface that is placed onto an object, or a thickness of some material laid on or spread over a surface. In many graphics editing programs, the working area is conceived as a set of layers, where higher layers' contents mask lower layers' ones. In Computer Science, an abstraction layer is a way of hiding the implementation details of a particular set of functionalities.

In brief, a layer is something that covers by adding contents, and it can be removed if lower areas have to be investigated.

Starting from these examples, the concept of layer can be applied to music information, too. In fact, as intuitively shown in the previous section, music information is made of heterogeneous facets, whose degree of abstraction may range from a purely logical description in terms of symbols to the physical description of audio signals.

In our opinion, the different aspects music information is made of can be organized into the following layers:

- a *General* layer, which contains music-related metadata, i.e. catalogue information about the piece;
- a *Logic* layer, which addresses the logical description of score symbols (e.g. chords and rests);
- a *Structural* layer, which potentially identifies a number music objects and their mutual relationships;
- a *Notational* layer, which embraces the graphical representations of music, such as printed or hand-written scores;
- a *Performance* layer, which is devoted to computer-based descriptions and performances of music, by employing so-called performance languages;
- an *Audio* layer, which contains audio tracks and video recordings of the piece.

A similar layout has already been discussed in some previous works, such as (Haus & Longari, 2005; Lindsay & Kriechbaum, 1999; Steyn, 2002).

All the mentioned aspects may have both an analogue representation in the real world and a digital version for computers. In this context, we are interested in exploring only the latter case.

It is worth mentioning that not all layers are necessarily present for a given music piece. For instance, many Jazz pieces have no score, since they have been originated by extemporary improvisation; at most, their performance can be transcribed *a posteriori*. Similarly, many traditional songs or popular tunes are poor as regards their metadata. Moreover, for each layer descriptions could exist but they could be unavailable at the encoding time. Of course, the higher the number of available layers, the richer the music description.

So far, richness has been mentioned concerning the number of heterogeneous descriptions, in a certain sense the number of supported media types. But in addition each layer could contain many digital instances. For example, the *Audio* layer could contain several audio tracks, as well as the *Structural* layer could provide many different analyses for the same piece. A graphical example is shown in Figure 2.

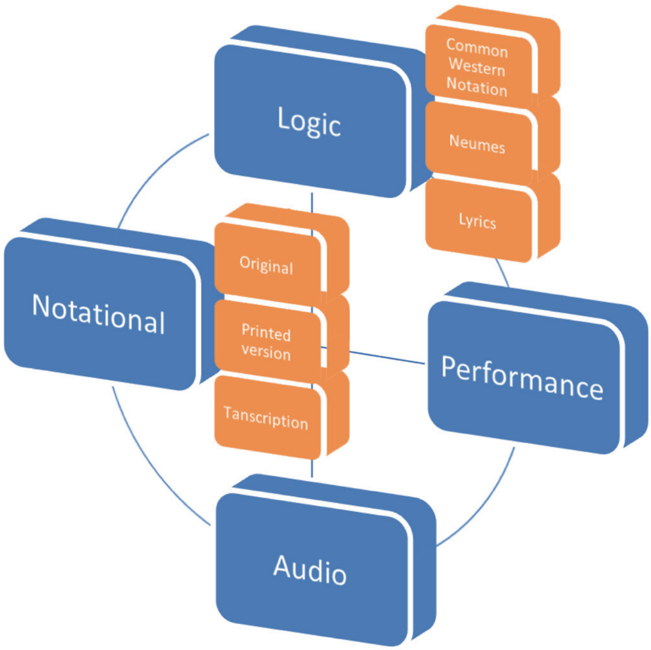


Fig. 2 - An example of multiple layers and multiple instances within single layers for music description.

The concept of multilayer description - i.e. as many different types of description as possible - together with the concept of multi-instance support - i.e. as many different media objects as possible for each layer - can provide a rich and flexible way to encode music in all its aspects.

3.2 Spatiotemporal Relationships

Both *space* and *time* are two key concepts for the representation of music. Their meaning and their function is fundamental in some of the mentioned layers, and this matter will be discussed in further detail in this section.

Temporal aspects of music are strictly related to music language itself. Time can be expressed both in absolute and in relative format. In particular, *absolute time* is computed on the base of an external clock and is measured in seconds, multiples and submultiples. This kind of description is commonly accepted for most multimedia formats. For instance, the duration of an audio track is typically mapped onto absolute clock time in terms of minutes, seconds and milliseconds. Similarly, video devices adopt SMPTE timecode or similar standards to synchronize audio and video contents.

On the contrary, *relative time* is a way to express duration in function of reference events or metric levels. This concept clearly refers to rhythm: each note and rest has a duration specified by its rhythmic figure. This kind of duration is not absolute, since it does not define how many seconds the music event should last; rather, it is relative to the speed of the beats specified by the metronome, namely a device that maps beats onto absolute timing. Many file formats, in particular those addressing computer-based performances (e.g. Csound, MIDI, etc.), adopt a relative definition of time, and it is sufficient to change the metronome parameter to obtain performances stretched in time. For details about these formats, please refer to (Selfridge-Field, 1997).

Consequently both absolute and relative time are crucial in music, and a comprehensive approach to music description should take them all into account. Also space plays a key role in music language. For instance, the standard way to notate a score in Common Western Notation (skipping details about paging and other graphical issues) is a 2-dimensional one, where both the horizontal and the vertical dimension have a precise meaning. The former aspect takes into account time flowing (melody, rhythm), whereas the latter indicates simultaneity among score symbols (harmony).

Time-related and space-related descriptions of the same music piece cannot be independent, since they provide different points of view on a unique information entity. The matter is finding a way to encode spatiotemporal relationships among instances belonging to the same or to different layers in a multilayer environment. In order to solve the problem, we can use the concept of spine. The spine is a data structure consisting of a sorted list of music events. The description levels of the mentioned multilayer layout can refer only to the events listed by the spine, which functions as a sort of glue among layers. At this point, a clarifying example is called for. In a piece made of n music events, the spine would list n entries. Each event has only one logic definition (roughly corresponding to score symbols), it may appear in many different score versions and rendered in a number of audio tracks. These aspects can be described in the Logic, Notational, and Audio layers respectively.

Music events are not only listed in the spine, but also marked by unique identifiers, so that each spine event can be described:

- in 1 to n layers; e.g., in the Logic, Structural, and Performance layers;
- in 1 to n instances within the same layer; e.g., in three different audio clips mapped in the Audio layer;
- in 1 to n occurrences within the same instance; e.g., the notes in a song refrain that is performed 4 times (thus the same spine events are mapped 4 times in the Audio layer, at different timings).

In a section that investigates the concept of spatiotemporal relationships, a key concept to introduce is the one of synchronization. *Synchronization* can be defined as an adjustment that causes something to occur simultaneously. In other words, synchronization is the coordination of occurrences to operate in unison with respect to time. As explained before, our work extends this concept to space as well. Thanks to the spine, heterogeneous instances of music description can be mutually synchronized, both time-based and space-based. Moreover, this data structure creates synchronization among instances within a layer (*intra-layer synchronization*), and also synchronization among contents disposed in many layers (*inter-layer synchronization*).

However, it is worth underlining that not all the music descriptions listed in our multilayer approach are synchronizable. Analyzing the proposed layout layer by layer can clarify this subject.

Computer-driven and human performances, described in the *Performance* and the *Audio* layers respectively, present temporal aspects, as music perception follows a timeline. As a consequence, they contain synchronizable events. Also the contents of the *Notational* layer, where only spatial aspects are explicit (i.e. the occurrence of symbols in a given position), can be considered synchronizable. In fact, the space position of symbols refers to spine events, which have their own timing. Similarly, thanks to spine references it is possible to determine the spatial position of the music events played in an audio track over a given score. Likewise, structural contents can be considered synchronizable, as they refer to spine events as well.

The only exception is represented by the *General* layer, which mainly provides metadata and catalogue information. For the sake of clarity, the fact that *Rhapsody in Blue* was composed by G. Gershwin, or the libretto for W.A. Mozart's *Le nozze di Figaro* was written by L. Da Ponte, are not related to time or to the occurrence of music events in time or space. Consequently, the contents associated to a music piece through the *General* layer are non-synchronizable.

4 Multilayer Music-oriented Learning Objects

In our approach, music information is structured according to a multilayer schema, and the spine is used to reference various descriptions of the same music event in each layer, and in each instance inside layers. The presence of the spine, which acts as the common data structure for the multilayer environment, is fundamental both for synchronization and for navigation purposes. In fact, it is possible to jump from one representation of an event to another, either within the same or in another layer, by referring to their common *id*.

Thanks to the spine, the network of interconnections among layers is automatically produced. As a desirable side effect, also the production of new media materials to enrich the encoding gets benefit from such a framework. In fact, adding a new digital object implies finding the references of music events to spine only within the new instance, whereas inter- and intra-layer synchronization with previously encoded materials is automatically achieved. A multilayer structuring of information can result in the design and implementation of advanced interfaces to enjoy contents. This idea will be developed in the next sections in order to create music-oriented serious games.

After introducing the concept of multilayer description, we need to provide a definition of learning object, in order to couple these key ideas and then apply them to serious games. In the literature, the terms "learning object" and "educational object" are used interchangeably. In writing this article, we adopt the former locution.

An in-depth discussion about learning objects can be found in (Sosteric & Hesemeier, 2002), where heterogeneous definitions are compared and critically analyzed. At its most basic level, a learning object can be seen as a piece of content that is smaller than a course. Unfortunately, such a definition is too general to be of any use in our discussion. In this sense, a further elaboration of the concept is provided by the IEEE Learning Technology Standards Committee (LTSC), as reported by (Friesen, 2001): a learning object is defined as any entity, digital or non-digital, which can be used, re-used or referenced during technology supported learning. The LTSC also provides a number of examples of these objects, including multimedia content, instructional content, learning objectives, instructional software and software tools, as well as persons, organizations, or events referenced during technology supported learning. To paraphrase the above definition, a learning object is anything that can be used during technology-supported learning in an educational environment. This sentence highlights the importance of context: in fact, not any physical nor digital object are automatically learning objects for the mere reason that they exist, but all them can become learning objects in a given context.

Recalling the previous sections, the multilayer approach to music description automatically generates learning objects, since both contents and their mutual relationships are embedded in a single information entity. In particular, the mentioned relationships make the context emerge: heterogeneous descriptions are not a collection of isolated contents, as they are interconnected within a common environment.

Narrowing the field to a digital workspace, a commonly acceptable definition is the one finally proposed in (Sosteric & Hesemeier, 2002): a learning object is a digital file intended to be used for pedagogical purposes, which includes, either internally or via association, suggestions on the appropriate context within which to utilize the object. In this sense, we can think to a learning object as a file that contains organized education-oriented information and a number of relationships among contents that constitute the context.

As regards the digital encoding of music-related information, there are many existing file formats aimed at music description. For example, AAC, MP3 and PCM formats are commonly used to encode audio recordings, GIF, JPEG and TIFF files can host digitized graphical scores, MIDI is a well-known format for computer-driven performances, etc. However, finding a unique file format able to catch all the different facets of music information and to organize them in a single environment is not trivial.

The recent international standard known as IEEE 1599 can be the answer: it implements both the aforementioned 6-layers structure and the spine concept, as discussed in (Ludovico, 2008). The format adopts XML to encode metadata, symbolic contents and the spine, whereas media documents are only linked and synchronized inside the XML document. The digital files referenced by an IEEE 1599 are not embedded, and they preserve their original format. IEEE 1599 has been applied to many fields, ranging from education and cultural heritage to advanced multimedia enjoyment.

A similar approach is followed by MPEG-7 working group on music (Bellini *et al.*, 2005). The aim is integrating symbolic music representation in the versatile multimedia framework provided by other MPEG domains. In this case, music does not constitute the information core to describe, but it is one of the possible domains supported by the format.

5 IEEE 1599-based Serious Games

An IEEE 1599 document is potentially very rich in information. Consequently, the characteristics illustrated before let us design a wide range of applications. For example, the matter of evolved music enjoyment for entertainment has been addressed by (Baggi *et al.*, 2005), whereas the valorization of music as a lively cultural heritage through IEEE 1599 has been illustrated in (Baratè *et al.*, 2011).

A field where IEEE 1599 can find application is music education and training. In fact, such a format supports a number of features that can be employed to create *ad hoc* implementations, e.g. a guide to music listening or a tool for instrumental and ear training. The heterogeneity of both descriptions and instances within each layer opens up new ways to enjoy a music piece. The matter to face here is how to implement these educational purposes in a computer game.

Before showing a serious game oriented to music education, let us deepen some concepts of the IEEE 1599 standard that will be extensively used in these applications. First, a "logical" score, i.e. the information encoded within the Logic layer in terms of music symbols, can be rendered in different ways:

- It can be dynamically reconstructed by a viewer/editor starting from the XML document itself, in particular by parsing the *Logic* layer;
- It can be linked to the *Notational* layer, which could contain not only printed or even hand-written scores, but also other forms of graphical description. This feature supports even scores not belonging to Common Western Notation.

Usually, a music piece has only one logic description. However, revisions, transcriptions for different ensembles, piano reductions, etc. can provide counter-examples relevant both from the cultural heritage perspective and for the design of music-oriented music games.

The *Audio* layer can host various performances of the same piece, which in general correspond to many different interpretations. From this point of view, cultural heritage are strongly involved. For instance, it is possible to encode, mutually synchronize and finally compare great historical recordings, e.g. Maria Callas and Renata Tebaldi performing the role of Cio-Cio-San during different seasons at the Teatro alla Scala. But - once again - the purpose of this layer can be extended and somehow forced. For instance, a multiplicity of tracks can be used to encode cover versions of same the piece, which can substantially differ from the original one. It sounds natural to apply this concept to jazz music and improvisation. Another possibility consists in using different audio instances representing the single tracks of a multi-track recording. Some layers, intentionally ignored until now, may allow many other applications. This is the case of the Structural layer, which permits the identification of music objects and their relationships in a score. The possibilities of the format are wide enough to embrace harmonic grids, segmentation, and different kinds of musicological analysis, as shown in (Dalmonte & Spaminato, 2008).

All the mentioned aspects will be used in the design and implementation of the following serious game, which is only an example to clarify the applicability of our approach.

[Score]FollowingPuccini, the serious game proposed in this section, specifically addresses musicians and students. The learning object concerns how to read and listen to music together, namely how to perform score following. Score following is the process by which either a musician or a computer system tracks the performance of a music piece. This tracking is done by following the progress of the music symbols in the score (written music) while the notes are playing.

This process is very important for music students to get knowledge and abilities about solfege, harmony, composition processes, music forms, and instruments' timbre recognition. The interface proposed in Figure 3 shows a handwritten music score and a performance by Chiara Panacci. Thanks to the multilayer approach, it is possible to switch the current score with a printed one, or to substitute the performer with another soprano.

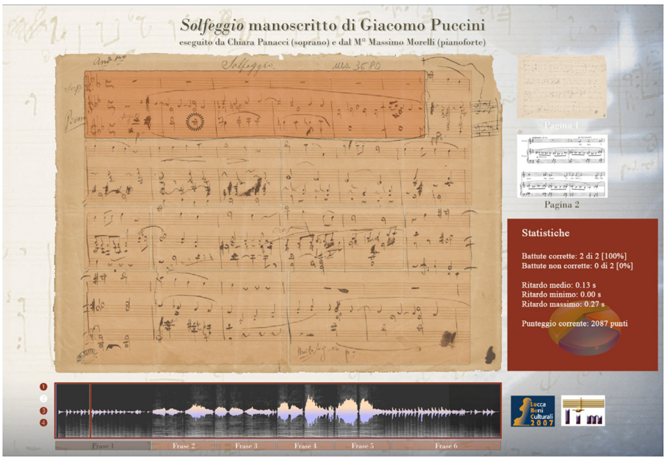


Fig. 3 - The interface of [Score]FollowingPuccini.

There are many stages in this game, corresponding to different score following activities. The most basic step, illustrated in Figure 3, is listening to music and highlighting the notes currently playing on the score through a cursor. The computer system detects misalignments and evaluates the corresponding offsets by assigning a game score. Another stage implies the inverse operation, namely identifying portions of the waveform by following the score, which is highlighted on the base of a given metronome. Finally, the most challenging stage of the game consists in detecting randomly selected parts of an audio performance on the score. The game becomes more and more difficult as audio segments becomes increasingly shorter and score areas smaller.

Conclusions

This paper has provided a number of definitions and examples addressing the application of serious games to music education. After showing a general approach to this issue, the second part of the work has focused a multilayer approach to music information, in order to get a rich and comprehensive description for a music piece. Thanks to its characteristics, the adoption of the IEEE 1599 standard enables a number of possible serious games oriented to music education. The case study described in Section 5 is just one of the many examples that can constitute the new frontiers in music education.

This work is one of the scientific outcomes of *Project E2*, funded by the Education, Audiovisual and Culture Executive Agency of the European Commission in the framework of the *Lifelong Learning Programme Leonardo Da Vinci - Development of Innovation* (E2 - 517964-LLP-2011-IT-LEONARDO-LMP). The authors wish to acknowledge all the researchers and contributors who made this result possible.

REFERENCES

Abt, C. (1970), *Serious games*. Viking Press.

Addis, M. (2005), *New technologies and cultural consumption–edutainment is born!* European Journal of Marketing 39, 7/8, 729–736.

Antonova, A. and Stefanov, K. (2011), *Applied cognitive task analysis in the context of serious games development*. In Third International Conference on Software, Services and Semantic Technologies S3T 2011, 175–182. Springer.

Baggi, D., Baratè, A., Haus, G., and Ludovico, L.A.(2005), *A Computer Tool to Enjoy and Understand Music*. In The 2nd European Workshop on the Integration of Knowledge, Semantics and Digital Media Technology (EWIMT 2005). IET, 213–217.

Baratè, A., Haus, G., Ludovico, L. A., and Mauro, D.A. (2011), *A Web-oriented Multilayer Model to Interact with Theatrical Performances*. In International Workshop on Multimedia for Cultural Heritage (MM4CH 2011), Modena, Italy. Communications in Computer and Information Science Series, vol. 247. Springer.

Bellini, P., Nesi, P., and Zoia, G. (2005), *Symbolic music representation in mpeg*.

Multimedia, IEEE 12, 4, 42–49.
Dalmonte, R. and Spaminato, F. (2008), *Il Nuovo in Musica: Estetiche, Tecnologie, Linguaggi*. Atti Del Convegno, Trento, 18-20 Gennaio 2008. Libreria Musicale Italiana.
Friesen, N. (2001), *What are educational objects?* Interactive learning environments 9, 3, 219–230.
Haus, G. and Longari, M. (2005), *A multi-layered, timebased music description approach based on XML*. Computer Music Journal 29, 1, 70–85.
Kurtzman, J. (1994), *Tones, modes, clefs and pitch in roman cyclic magnificats of the 16th century*. Early music 22, 4, 641–665.
Lindsay, A. and Kriechbaum, W. (1999), *There's more than one way to hear it: multiple representations of music in MPEG-7*. Journal of New Music Research 28, 4, 364–372.
Lorenz, K. and Kickert, R.W. (1987), *The waning of humaneness*. Little, Brown.
Ludovico, L.A. (2008), *Key concepts of the IEEE 1599 standard*. In Proceedings of the IEEE CS Conference The Use of Symbols To Represent Music And Multimedia Objects, IEEE CS, Lugano, Switzerland. 15–26.
Parrish, C. (1978), *The notation of medieval music*. Pendragon Press.
Ramstetter, C. L., Murray, R., and Garner, A.S. (2010), *The crucial role of recess in schools*. In Journal of School Health 80, 11, 517–526. Wiley Online Library.
Read, G. (1979), *Music notation: A manual of modern practice*. Taplinger Publishing Company.
Rubsamen, W. (1968), *The earliest French lute tablature*. Journal of the American Musicological Society 21, 3, 286–299.
Selfridge-Field E. (1997), *Beyond MIDI: The Handbook of Musical Codes*. The MIT Press, ISBN: 0262193949
Shirali, V. (1977), *Sargam: an introduction to Indian music*. Abhinav/Marg Publications.
Sosteric, M. and Hesemeier, S. (2002), *When is a learning object not an object: A first step towards a theory of learning objects*. The International Review of Research in Open and Distance Learning 3, 2, Article–3.
Steyn, J. (2002), *Framework for a music markup language*. In Proceeding of the First International IEEE Conference on Musical Application using XML (MAX2002). 22–29.
Union, W. B. and Krolick, B. (1996), *New international manual of Braille music notation*. SVB.
Zyda, M. (2005), *From visual simulation to virtual reality to games*. Computer 38, 9, 25–32.