Abstract
This paper begins with some of the topics and questions about music improvisation using computers, the author has posed himself during more than a decade experience as a performer, composer, software designer and educator. After a quick review of several of his previous interactive music systems, some of those issues are then confronted in FMOL, a program initially designed as a tool for on-line collaborative composition that has been used by hundreds of on-line composers, and which is also being employed by the author in free-form audiovisual improvisation concerts.

1 Introduction
Improvisation is today present in almost every area of music, although it may have completely different meanings, behaviors and rules depending on the idiom it applies to (e.g. baroque, bebop, flamenco, blues, rock, etc.) and can even attempt at being non-idiomatic as in the so-called free improvisation (Bailey, 1992). Considering this infinite variety of approaches, further categorizations, generalizations and even definitions become really hard if not impossible tasks; we will try to stick therefore to what could be the common denominator of all these practices, and consider improvisation as any form of instant composition, no matter how free or constrained this compositional process may result.

Computer-based interactive music systems go back to the late 1960s, initially involving computer-controlled analog synthesizers in concerts or installations. The use of real-time algorithmic composition spread in the 1970s with the work of composers and performers such as David Behrman, Joel Chadabe, Salvatore Martirano, Gordon Mumma or Laurie Spiegel (Mumma, 1975; Bernardini, 1986), but its greatest impulse really came during the mid 1980s with the MIDI standardization and, slightly later, with the advent of data-flow graphical programming languages such as MAX (Puckette, 1988; Puckette and Zicarelli, 1990), which made the design and implementation of custom interactive systems simpler than ever before (Winkler, 1998).

However, one and a half decades later, improvisation using computers still seems a burgeoning and under explored multidisciplinary area where the design of new controller interfaces, real-time sound synthesis and processing techniques, music theory, cognitive science, algorithmic composition techniques, and existing models of improvisation (computer and non-computer based) can converge, in order to bring new interactive music-making paradigms. If computer music improvisation must try to define new models, it obviously cannot ignore the existing ones. Computer music improvisers have therefore in front of them a myriad of practices and know-how’s, several as old and rich as the history of music making, and also a half-century-old tradition, that of computer music.

The next part of this paper introduces three significant topics (others could have been included), discusses their current situation related to either instrumental or computer music, and suggests how fresh approaches can surpass traditional limitations and influence in the conception and implementation of new improvisation models.

2 In Search of New Computer Music Improvisation Paradigms: Three Topics to Think About

2.1 Controllers and Generators
Real-time interactive music needs musical input devices (i.e. controllers). The separation between gesture controllers and sound generators standardized by MIDI boosted the creation and development of many new alternative controllers with which to explore new creative possibilities (Chadabe, 1997). However, this separation should not always be seen as a virtue or an unavoidable technical imposition.

Before MIDI, the term musical instrument always referred to both sides of the sound and music creation process (i.e. controller and generator), and it is obvious that the final expressiveness and richness of any musical interface (controller) is not independent of the generator involved and the mapping applied between them. Some of the consequences that result from this fact are truisms (e.g. not all combinations of controllers-generators are equally satisfying), but there are also some important questions that raise: is it possible to develop highly sophisticated and efficient controllers without a prior knowledge of how the
sound or music generators attached to them will work? We firmly believe that it is not possible and that a parallel design is necessary. Are today’s standard music communication protocols (i.e. MIDI) wide and flexible enough, or are the potential dialogues between controllers and generators limited or narrowed by these existing protocols? Although MIDI has many limitations (low bandwidth, low resolution, half-duplex communications, etc.) (Moore, 1988) we optimistically believe that there is still room for experimentation and improvements while keeping the compatibility with the MIDI constraints.

Bi-directional mappings may be one of the solutions leading to wider and richer communications between controllers and generators and, as a result, between players and their music output. Duplex communications allow the existence of haptic and force manipulations and feedback, which may bring richer feedback loops between the instrument and the performer (Gillespie, 1999).

But force feedback and by extension, performer-instrument feedback, are not the only feedback loops present in traditional music playing: acoustical instruments, for example, resonate (i.e. they are ‘conscious’ of the sound they produce) and this acoustic feedback that exists between the controller and the generator subsystems, is in fact fundamental in the overall sound production process. While audio analysis, which could be extended to the concepts of music analysis or machine-listening, is a regular topic in many computer-music disciplines, this task is not usually undertaken by the controller subsystem, meaning that too often, controllers are not ‘knowing’ the output they are generating.

With the search for smarter mappings and wider communications between the different components of an interactive music system, we are not implying however that low-cost and widely available input devices such as mice, joysticks or computer keyboards have to be considered as an all-gone dead-end. On the contrary, there is still a lot of research to be done in the area of interactive GUIs. We will see, how in the case of the mouse-driven FMOL instrument, the controller audio feedback presented to the performer in a visual form, intuitively helps the understanding and the mastery of the interface, enabling the simultaneous control of a high number of parameters that could not be possible without this visual feedback. This special feature converts FMOL into an audiovisual synthesizer.

2.2 Macro vs. Micro Level Control

For centuries, Western music has talked about ‘notes’, almost ignoring what was inside of them. However, many instrumental and vocal improvisation idioms keep actually a good balance between sound and notes —or form— (consider for instance the growls of free jazz saxophone players or the melismas of flamenco singers, both essential to their respective improvisational styles).

While it is true that since its beginning, half century ago, computer music started considering and giving a relevance to the notes' inside, the prevalent use of the two complementary concepts ‘score’ and ‘orchestra’, common to most Music N languages, did certainly not encourage the merger of these two, macro and micro levels. But nowadays, the fact is that the separation of both worlds, when applied to computer music creation, may be becoming not only less and less trivial, but can also be seen a burden and an anachronism.

With the advent of MIDI, almost two decades ago, computer assisted music creation (as opposed to computer assisted sound creation) was able to become a real-time activity. MIDI does indeed allow some sound parameters modifications by means of control change messages, in degrees that vary with each synthesizer model, but its paradigm is completely note-based. The panorama may have started changing these very last years, when the increasing power of personal computers has allowed the definitive bloom of real-time software synthesizers. Even though virtual synthesizers do not really represent any new concept, they have important differences with their hardware counterparts, as they have the potential to achieve what no manufacturer could ever think of, giving more room for freedom, experimentation and imagination, and bringing forth new sonic and control paradigms. Sadly, control over most of these software tools, still sticks to the idea of notes which can have an internal evolution, thus keeping the tacit computer-music distinction between instruments (that control sound) and composing tools (that control form).

This dichotomy is also implicitly present in the approach of most improvisers and interactive systems designers nowadays. The more composition-oriented musicians tend to the macro-level idea and others favor the more performance or instrument-oriented micro-level one, while a third group may try to find an uncertain equilibrium between both conceptions.

In this context, several questions arise: wouldn’t it be possible to conceive performing systems (controllers + generators) where the control of both sound and form could be done by means of the same tools and gestures? And wouldn’t that bring a completely new dimension to real-time computer music creation?

Several attempts have been made in that direction. For instance, the Music Mouse interactive software, developed by Laurie Spiegel in the mid 1980s, can be considered both an instrument and a composing tool (Gagne, 1993), but the technology available fifteen years ago was not mature enough for this kind of mixed approaches. My background as an amateur free jazz saxophonist in the early 1980s, may surely be one of the reasons for which I seek in this no man’s land territory. And this is also one of the fundamental aspects FMOL tries to explore.
2.3 Individual vs. Collective and Dilettante vs. Professional Performers

Creating music in real-time using computers and new interfaces poses also new questions about whom and how will be using them.

While performing music has typically been a collective event, traditional musical instruments have been mostly designed for an individual use (even if some, as the piano or the drum kit can be easily used collectively). This restriction can now be freely avoided when designing new interfaces, which leads to a new generation of distributed instruments, with a plethora of possible different models following this paradigm (statistical control, equally-allowed or role-playing performers, etc.). Implementations of musical computer networks date back to the late 1970s with performances by groups like the League of Automatic Music Composers or the Hub (Bischoff, Gold and Horton, 1978). This may also be the common case of many interactive sound installations which respond to the public movements and gestures, which leads us to another important point: that of the skills and the know-how of the performer(s).

I have developed several computer-based interactive music systems since 1989. Some of them were conceived for trained musicians or even for specific performers, while others, like Epizoo, were to be controlled by members of an audience in public performances (Jordà, 1996; Lozano-Hemmer, 1996). The demands for the two genres are obviously different. Complicated tools, which offer great freedom, can be built for the first group, while the second group demands simple but appealing tools that -while giving their users the feeling of control and interaction- must produce 'satisfactory' outputs. These two classes are often mutually exclusive. Musicians become easily bored with the 'popular' tool, while the casual user may get lost with the sophisticated one. But is this trend compulsory? Isn't it possible to design interfaces that can appeal to both sectors: tools that would not dishearten hobbyist musicians, but that would still be able to produce completely different musics, allowing a rich and intricate control and offering various stages of training and different learning curves? Tools that, like many traditional acoustical instruments, could offer a low entry fee with no ceiling on virtuosity (Wessel, Right, 2001).

There is the common belief that more hard-to-play instruments lead to richer and more sophisticated musics (e.g. the piano vs. the kazoo), but expressiveness does not really imply difficulty, and in that sense, one of the obvious research trends in real-time musical instruments design can be the creation of easy-to-use and, at the same time, sophisticated and expressive systems. Considering that the best way to understand and appreciate any discipline, whether artistic or not, and music is no exception, is by doing and being part of it, more efficient instruments that can subvert the previous effort-result statement will bring new sophisticated possibilities and the joy of real-time active music creation to non-trained musicians. Lets try, as Robert Rowe suggests, to develop virtual instruments and/or musicians that do not just play back music for people, but become increasingly adept at making new and engaging music with people, at all levels of technical proficiency (Rowe, 1992).

3 Three Previous Works by the Author

3.1 PITEL

Influenced by the works and writings of George Lewis (Roads, 1985) and Robert Rowe (Rowe, 1992), PITEL (1989-1992) is the first interactive music system I developed (Jordà, 1991). It is a software environment for polyphonic real-time composition and improvisation, which employs some simple mathematical machine-listening techniques, that can generate, under the control of a mouse-conductor, up to four monophonic MIDI voices, while listening and reacting to one or two external MIDI players.

PITEL performances were usually jam sessions that involved three musicians: a saxophone and a trumpet player, both fitted with pitch-to-MIDI converters, and myself conducting the program. All high-level compositional parameters were adjusted by simple means of virtual buttons and sliders. The aesthetical model was closer to free music than to the jazz tradition (i.e. no scales, modes or predefined rhythms were used), but the instrument remained at the note and form level, without attempting any sonic micro-level control as mentioned in 2.2.

I wanted PITEL to be as flexible as possible and not oriented to any special musical aesthetic, so I decided to build a system with no musical knowledge at all (except that one octave has 12 MIDI notes), as I felt that introducing some musical ideas into the system, would have favoured particular musical styles.

![Figure 1. PITEL listening matrix. In this configuration, all four internal voices are listening to voice #2 (although using different algorithms), creating a slightly homophonic effect. Voices A and B correspond to the human players.](image)

The program was based on two-term non-linear feedback relations of the form \( X_{a,i} = F(X_{a,i-N} , X_{b,i-k}) \) (where \( a \) and \( b \)
indicate two different musical voices). The output of any voice could then be influenced by its own past and by the past of another voice. The matrix shown in figure 1, allowed to configure in real-time to what voice (including the two human ones represented by A and B) was each one of the four PITEL voices listening to, and which listening algorithm it was using.

This simple approach produced remarkably coherent music between the four internal voices, but the results were not so satisfactory when these internal voices were listening to the human players, as they were not sharing the same listening logic (or algorithm). Two of the other criticisms that could be applied to the system were its big inertia (i.e. fast changes were hard to obtain) and, as shown in figure 2, the shortage of a serious interface design and conception (which is not necessarily the case of all mouse-driven GUI).

Figure 2. PITEL Main Screen.

### 3.2 The QWERTYCaster

While PITEL algorithmic and inertial approach follows perfectly the composition tool paradigm, the low-tech QWERTYCaster (1996) is a clear example of the instrument model. It is a very simple guitar-shaped controller that I designed solely for myself and used for free improvisation. The aim was to build a fast system, with direct audio output in response to every movement or gesture, not unlike Nicolas Collins 'Trombone' (Collins, 1991) or Michael Waisvisz 'Hands' (Krefeld, 1990). It consists of a QWERTY computer keyboard (strings), a trackball (frets) and a joystick pad (lever), all held together in a guitar-like piece of wood, connected to the computer with a five-meter cable which allows me to play the 'rock guitar hero' role. Its five continuous controllers (two degrees of freedom for the trackball and three for the joystick), together with the joystick triggers, buttons and keys, steered an old 486 computer with a sampler soundcard and a simple but effective custom MIDI software. At the opposite side of PITEL, the QWERTYCaster only focused in notes and sound control, leaving all the macro-formal organization for the human improviser. In spite (or maybe because) of its simplicity, I still use it sometimes in improvisation concerts, especially in duos.

![Figure 3. Sergi Jordà with the low-tech QWERTYCaster (1996).](image)

### 3.3 Afasia

Afasia (1998), my third and last collaboration with the visual artist and performer Marcel.li Antúnez (Lozano-Hemmer, 1996), is a one-man-show multimedia play inspired in Homer's Odyssey, which has been performed in many countries in Europe and America and has received international awards. In Afasia, the performer-conductor (Antúnez), fitted with a plethora of radio-emitting sensors, controls with his movements, a whole mechanical music robot quartet (consisting on a seventy-fingers electric guitar+bass, a walking drum kit, a one-string electric violin and a three-bagpipe 'horn section'), a sampler, an audio CD, three audio effects racks, a Yamaha Promix MIDI-controlled audio mixer, the interactive multimedia animations, a DVD, a DMX light table and the video projector input-switcher (for changes between SVGA multimedia and DVD).

All of the sonic elements of the system are interactively MIDI controlled. Each one of the four robots has its own virtual MIDI driver which bridges musical MIDI messages generated by the interactive software, with the digital output cards that control the relays and the pneumatic mechanisms. The music mechanically played by the electric guitar+bass and the violin robots (the drum kit robot and the three bagpipes are not amplified), is mixed together with the music coming from an internal sampler card and with some prerecorded audio fragments triggered from an internal CD Audio player, into the Yamaha mixer, and this output is
eventually processed through three multieffects processing units, which are also MIDI controlled.

The complexity of this setup as opposed to the limited semantics of the sensors employed (gloves, buttons, potentiometers in each of the performer’s articulation and mercury switches in his extremities) makes Afasia a peculiar example of a score-driven interactive model, in which every ‘island’ (taken from Ulysses’ journey) behaves as an independent interactive environment, which allows the performer to interact with different preset mappings and several restricted degrees of freedom.

Figure 4. Marcel.lí Antúnez controlling the three-bagpipes robot in Afasia (1998).

3.4 Afasia Interactive MIDI Files

These previous setup considerations lead me to implement Afasia’s interactive music capabilities (images and video are taken differently) by means of format 1 Standard MIDI Files, expanded with the use of text meta-events.

Text events are often used by sequencers to name parts of the tracks or display synchronized lyrics. In the Afasia software they are used as commands that tell the custom sequencer player how to deal with, and process the recorded events, according to the performer’s gestures. This approach enables to write, record and program interactive Afasia music using any standard sequencer software, although the interactivity is only available when the file is played in the custom Afasia sequencer player (when Afasia Standard MIDI files are reproduced on a conventional sequencer no interactivity is attained, as sequencers simply skip text meta-events). Here is a brief description of the Afasia MIDI files structure:

- Each part (or island) of the show is associated with one Standard MIDI file.
- Each MIDI file is made of any number of blocks.
- Blocks (an Afasia concept not present in standard MIDI files) are like permutable sections of a score, and only one block can be active at any time.
- Each block is made of a special control track and any number of conventional MIDI tracks.
- Control tracks only contain text meta-events that indicate how to interact with the block’s tracks, alter the block structure or jump to other blocks.
- Conventional tracks can also contain text meta-events that indicate how to interact or modify the MIDI data inside the track.
- As conventional MIDI tracks, each track is directed to a specific port or device (one of the four robots, the internal sampler, the audio mixer or any of the three effects processors), and to a specific channel within that device.
- Each device can have several MIDI channels (e.g. the electric guitar+bass robot has one channel for each of its strings).

Text meta-events embedded in the MIDI tracks, specify interactive commands and have always the same structure:

```
COMMAND_NAME=param1,param2,...,paramN$
```

The number of parameters and their meanings varies with each command type. The first parameter usually corresponds to a sensor number while the second indicates the track of the block the command applies to. Following parameters define how these incoming sensor values have to be mapped. The command \( \text{TRANSPOSE}=4\,3\,12\) indicates for example that the third track of the block will be transposed according to the values of sensor #4, up to a maximum value of \(+12\) semitones. Many commands (TRANSPOSE is one of them) have an ‘end’ version that disables the effect.

More than forty commands are implemented in Afasia. They allow to dynamically switch between blocks, mute/unmute tracks, apply loops and modify their lengths, modify the current play position, generate or apply changes to any MIDI parameter (controlling its new value, its increment, or the variation random range), quantize tracks, delay tracks, control tempo, compress/expand events, define chords or scales (to which generated notes will be filtered), etc.

Although I have not performed with the system myself, I have composed several interactive pieces for the show, to be performed by Marcel.lí Antúnez. The system has proved to be a flexible framework for interactive music composition and performance and the model seems also especially suitable for interactive soundtracks in videogames or multimedia CDROMs.
4 FMOL (F@ust Music On-Line)

The FMOL project (1997-2001) started when the Catalan theatre group La Fura dels Baus, proposed me to develop an Internet-based music composition system that could allow cybercomposers to participate in the creation of the music for their next show, F@ust 3.0. Initially I did not have a clear idea of what I wanted; I just knew what I did not want: to allow people to compose tempered music on the keyboard and send us attached MIDI files via E-mail. Besides, although I felt that the project should have a fairly 'popular' approach, and did not want to be too demanding and restrictive about the participants' gear, I was not looking for a dull General MIDI sound, but for richer sounds and textures that could introduce newcomers into more experimental electronic music (Jordà, 1999). Mouse-driven software for real-time interaction and synthesis seemed therefore the natural solution, which led to the development of a C++ stand-alone program with http capabilities (Jordà and Wüst, 2001).

4.1 FMOL's Interface Design

The conception of Bamboo, FMOL's main graphical mouse-controlled interface, was very tight with the synthesis engine architecture design. Both were in fact developed in parallel with the primary aims of conceiving a real-time composition and synthesis system, appealing to both trained electronic composers and more casual or hobbyist musicians, suitable for the Internet (small scorefiles, etc.), that could run on a standard computer fitted with any conventional multimedia soundcard.

Discussing FMOL's architecture would be too long. I will just mention that the sound engine supports eight stereo real-time synthesized audio tracks, each one consisting of a sound generator (sine, square, karplus-strong, sample player, etc.) and three serial processors (filters, reverbs, resonators, frequency or ring modulators, etc.), which can be chosen by each composer between more than 100 different synthesis algorithms or variations. Each generator or processor contains in its turn, eight control parameters, four of which can be modulated by four independent low frequency oscillator (LFOs), in which frequency, range and shape (sinusoidal, square, triangular, saw tooth or random) can also be interactively modified for each. The important point is that the whole control of this complicated architecture is clearly reflected in an intuitive, symbolic, dynamic and non-technical way, in Bamboo's graphical interface (Jordà and Aguilar, 1998).

The Bamboo screen, presents a lattice in which vertical lines are associated with the synthesis generators and horizontal lines with the synthesis processors. Like a virtual guitar, these vertical lines/strings can be plucked or fretted with the mouse while they continuously draw the sound they generate like a multichannel oscilloscope. Horizontal segments, on the other hand, control the synthesizer's serial processors, and can be dragged and oscillate up and down, creating an abstract dance that tightly reflects the activity of the piece.

A Bamboo user's manual cannot be included but I will just mention that the combination of both mouse buttons and the computer keyboard allows for an intricate control, including sustaining sounds, modifying any parameter, recording mouse gestures loops, applying LFOs (with frequency, amplitude and wave-shape control), creating arpeggios or recording and restoring sequences and screen configuration snapshots. Although it may sound quite complex, the abstract visual feedback of all the instrument activity has proved to be an invaluable help for users.

Figure 5. FMOL’s Bamboo in full action

4.2 Low Frequency Oscillators and Time Control in FMOL

The FMOL software can act as a sequencer and as a player, and this facility is indeed fundamental for submitting, retrieving and playing compositions (scorefiles) via the web. However, for conceptual and esthetical reasons, record and playback capabilities were somehow restricted in order to avoid the triggering of pre-recorded sequences while improvising. Instead of that, users can dynamically record and retrieve all their mouse movements, as many times as they want with the help of two keyboard keys, but this material vanishes when s/he leaves the session. Other possibilities include the storage and retrieval of snapshots that hold the state of all the parameters at a given time, and all these snapshots may be indeed stored with the scorefile. Although snapshots are instantaneous (they could be called presets using a synthesizer terminology), they do include the information for the twenty-four active LFOs (frequency, amplitude and wave shape), and the combination and interaction of all these LFOs are in fact responsible for much of the time control and the dynamic evolution of
FMOL pieces. In that sense, FMOL is an attempt at integrating real-time composition with simultaneous micro and macro control, in one only interface.

4.3 Musical and Social Implications

Unlike any other system I have designed, FMOL has been used by hundreds of Internet composers. From January to April 1998, the FMOL first Internet-database received more than 1,100 brief pieces by around 100 composers, some of whom connected nightly and spent several hours a week creating music. One of our main goals (i.e. to conceive a musical system which could be attractive to both trained and untrained electronic musicians) was fully attained. We know now that several of the participants had no prior contact with experimental electronic music and a few were even composing or playing for the first time, but all of them took it, however, as a rather serious game, and the final quality level of the contributions was impressive. After a difficult selection process (only 50 short pieces could be chosen and included on the show's soundtrack), and considering that a great number of interesting compositions had to be left aside, we decided some months later to produce a collective CD with a mixture of old and new compositions.

A new web with a new version of the software has been back on-line during September 2000 for La Fura’s new show, the opera DQ, premiered last October at the Gran Teatre del Liceu in Barcelona. During one month, more than 600 compositions have been submitted, and the selected ones constitute now the electroacoustic parts of an otherwise orchestral score.

In September 2000, a one-week workshop specially aimed at visual artists with no prior musical knowledge, took place in Lisbon. The workshop concluded with several conducted collective improvisations, with astonishing results.

4.4 The FMOL Trio

Although FMOL was originally designed as a freely available system for on-line collaborative composition and 'experimental electronic music proselytism', it also turned to be my favorite instrument for live concerts. Since 1999, the FMOL Trio (Cristina Casanova and myself on FMOL computers, plus Pelayo Arrizabalaga on saxophones/clarinet and turntables) performs free-form improvised electronic music, while two projectors connected to each of the computers give the complementary visual feedback. This setup, which enables the audience to watch the music and how it is being constructed, has proven to be a valuable addition to the concerts, giving the public a deeper understanding of the ongoing musical processes and adding new exciting elements to the show.

5 Conclusion and Future Work

After years designing and performing with interactive music systems, my favorite one has become the one I had not explicitly designed for myself, but for anyone. For the last four years, FMOL has become my main research and production area, and further developments of the FMOL instrument have taken different courses: as a model for Internet collective music creation, as a free software musical instrument and as a professional system specially oriented to our own performances.

The new Internet version, which should be available at the beginning of 2002, will finally feature real-time interaction between several users that will be able to share a common environment (i.e. they will improvise together) (Jordà, Wüst, 2001). This version may also include some additional and simpler game-like interfaces with restricted freedom, for the less adventurous composers. In parallel, we are developing a special concert version, which will include sensor-gloves, and will substitute the mouse with video detection, in order to allow the performers to play directly with their hands over a 3x2 meters retro-projected FMOL screen.

Figure 6. The FMOL Trio in performance. From left to right: Sergi Jordà (FMOL computer), Cristina Casanova (FMOL computer) and Pelayo F. Arrizabalaga (alto sax) (2000).
References


On-line Additional Information

Audio and video excerpts for this paper:
http://www.iua.upf.es/~sergi/download/chi2001/

FMOL main page: http://www.iua.upf.es/~sergi/FMOL

FMOL-DQ Internet composition project:
http://teatredigital.fib.upc.es/dq

FMOL Trio: http://www.iua.upf.es/~sergi/FMOL/fmoltrio/