ABSTRACT
This paper presents a brief review of current literature detailing some of the issues and trends in composition and performance with interactive music systems. Of particular interest is how musicians interact with a separate machine entity that exercises agency over the creative process. The use of real-world metaphors as a strategy for increasing audience engagement is also discussed.

The composition and system Church Belles is presented, analyzed and evaluated in terms of its architecture, how it relates to existing studies of musician-machine creative interaction and how the use of a real-world metaphor can promote audience perceptions of liveness. This develops previous NIME work by offering a detailed case study of the development process of both a system and a piece for popular, non-improvisational vocal/guitar music.

Author Keywords
Interactive music systems, real-world, metaphor, physical model, popular music

ACM Classification
H.5.5 Sound and Music Computing [Methodologies and techniques], J.5 Arts And Humanities [Performing arts (e.g., dance, music)]

1. BACKGROUND
1.1 Musician-Machine Interaction
In the field of interactive music systems, the idea that system building is part of the compositional process is well established [6]. Although by no means mainstream, many musicians conceive and construct their own software tools to act as partners with which they compose and perform [3]. Joel Chadabe discusses the process of “interactive composition” with such systems as a two-stage process: first the act of creating the system itself and then with playing it [7]. More recently, the benefits of emergent design through multiple iterations of this process have been observed [13].

The terms “memetic agency” and “performative agency” refer to the system’s influence over the composition in non-real-time and real-time [5]. Memetic agency may include serendipitous happy accidents that may result from idiosyncrasies or glitches in the system; or by the ongoing refinement and redevelopment of software. Performative agency refers to the influence of the novel, unpredictable musical output due to some kind of autonomy built into the system.

Many interactive music systems are used as duetting partners with human performers, where the machine listens to the signal emanating from the musician’s instrument, converts it into control data which is mapped to various system parameters. A plethora of signal analysis and machine learning tools enable meaningful performance data to be collected from instrumental/vocal output without needing to create complex, sensor based systems [16] or overload the human performer with additional tasks [2].

Andrew Johnston, Linda Candy and Ernest Edmonds report three modes of interaction of musicians playing with systems in real-time [16,17]. In the “instrumental” mode, the human performer attempts to control the machine performer. In “ornamental” mode, the machine is allowed to generate a separate accompaniment to the human output. In “conversational” mode, there is a switching between the two and therefore a shifting power balance in terms of agency. These findings fit with earlier classifications of interactive systems that are organized by notions of how much control is exerted, either by the human or by the machine performer [23].

1.2 Musician-Machine Performance
1.2.1 Liveness
The idea of musicians performing with autonomous machines is still fairly new to many audiences [3]. Recent NIME work has recognized a need to consider digital instrument and interaction design from the audience’s perspective, as well as from the performer’s [2,20]. Paul Sanden discusses several factors that influence an audience’s perception of liveness and therefore their engagement [21]. Of particular interest here are the ideas of presence (physically being there), corporeality (link between physical and sonic gestures), interactivity (on-stage interactions between musician and machine) and virtual liveness (the perceptual combining of human and machine agents as “performing cyborgs”).

Traditionally, the machine elements of the performance have been hidden from view. However, the potential for enhancing the power of a piece can be enhanced through the on-stage presence of both human and machine collaborators [24]. Strategies range from simply positioning system hardware onstage with displays clearly visible [2] to the use of humanoid robots [14].

It is now widely accepted audiences have a desire to understand how sounds are produced by machine performers [10,11], and how they are interacting with their human partners [20]. If this understanding is not established, a sense of frustration and disengagement may result [1]. Fostering this understanding is most commonly achieved by providing visual information. Perceiving a performing body is often equated with seeing it [21] despite evidence that audiences are aware of generative processes through listening alone [4]. However, one of the advantages of using visuals to represent these systems is that non-audible gestures can also be perceived, helping to communicate the intentions and interactions of performers [21].
Strategies for creating transparent processes include visually augmenting interactions [19], the use of virtual reality [24] and the use of real-world models that are visually represented on stage [18]. Skilled performers can also play in a way that highlights the interactive capabilities of their machine partner [3].

While the idea of “performing cyborgs” may result from instances where the performer is somehow merged with their system (e.g. through the use of wearable technology such as in Imogen Heap’s *Me the Machine* [15]), there are many instances where interacting musicians and machines are very much separate entities (e.g. Gordon Mumma’s *Hornpipe* [23]).

The next section details how this separation and interactivity can be enhanced through the use of physical models and real-world metaphors.

1.2.2 Real-World Models and Metaphors

Models based on real-world metaphors for interactive, generative machine performers have included models of masses on springs [16], dynamic fluids [18], rotating strings [13] and virtual objects in 3D virtual reality [24].

The use of real-world metaphors offers several advantages, particularly when working with a human performer. These include a shared mental model of the system between system designer, performer and audience [2] that represent “embodied understanding” [19].

Such systems facilitate simple, intuitive mappings of human performance data to system behavior – for example, a specific pitch can be mapped to a specific object in the physical system. The velocity of the note can be directly mapped to a force exerted on an object [16]. This circumvents the problem of processing large amounts of gestural input data into multiple parameters and devising complex mappings to the workings of generative algorithms [22]. The simplicity of the mappings afforded by the human-real world paradigm enable them to be manipulated during the course of a performance. This adds to the interest of the work whilst retaining the engagement of the audience [18,23].

A further advantage relates to the burden on the composer-programmer. Whilst widely viewed as an excellent way to generate highly interesting and varied results, artificial intelligence systems employing computational creativity are beyond the programming ability of many composer-performers [3]. Using real-world models avoids the need to incorporate artificial intelligence by endowing the machine collaborator with dynamic, Newtonian behavior. When interacting with a human performer, this is sufficient to yield interesting and complex results [13].

Interactive systems and pieces for human and machine performers that make use of real-world metaphors and co-performers include the *Partial Realities* series [16,17] that model masses and springs. *Encoded* [18] uses fluid dynamics as its central metaphor. Pieces such as *Virtual Real* [24] and *Carillon* [8] explore physical metaphors in mixed reality performances where performers interact with the system through its visuals. *Stringtrees* [13] is a mechanical system with generative elements that a performer manipulates directly.

1.3 Genres

The use of co-performing interactive systems has mostly been employed for free-improvisational styles involving classical performers [16], though some systems aim to appeal to broader audiences. *Algorave* [9] and *Virtual Real* [24] have been used to create and perform IDM, while *Genjam* [2] involves traditional jazz styles. To date, and as far as the author is aware, there has been little work done in this field involving the performance of popular music with live guitar and vocals.

2. CHURCH BELLES

2.1 Aims

The main compositional aim was to create an interactive, audio-visual piece of music combining chaotic rhythmic and melodic elements with traditional popular music structures and instrumentation. This author shares the view that pieces are most successful when they strike a balance between satisfying and challenging an audience [12].

The project was also an investigation into how a system based on a real-world metaphor could act as a compositional and performance partner, with both memetic and performative agency, able to demonstrate a high degree of liveness in performance settings.

2.2 Methodology

Similar to previous work in this field, this project is an example of practice-based research [13]. Detailed descriptions and reflections on the iterative activities of system building and playing with it were recorded as the composition emerged.

Although informal feedback from others has certainly contributed to development, formal audience evaluations have not been carried out, as these can often be of limited use [3]. Instead, the audience’s perspective was considered in the initial conceptualization and throughout development [20].

The system has not been tested with other musicians [16], as building it is considered to be part of the author’s compositional process for a particular piece. It is not intended that this system would be performed by other musicians or even be used for other pieces without making considerable adjustments.

Instead of more formalized evaluation, observations and insights from reflective practice will be compared with existing works and theories from the literature.

2.3 Design Criteria

*Church Belles* was designed using existing design criteria for similar human-machine systems [16]:

- The system should provide audio and visual responses in real-time to live audio from the human performer.
- There should be no need for additional sensors or interfaces other than the live audio from the performer.
- Expert musicians should be able to exercise control over the system.
- The audio and visual outputs of the system should be complex and engaging.
- The system and its interaction with the human performer should be transparent.

With the following additional criteria:

- The system should be fully interactive in that it directly affects the human performer’s output as well as being controlled by it [21].
- The system’s behavior and mappings should be able to develop during the performance of the piece [18,23].
- The visual element of the system should be kept simple in order that it does not detract from the musical aspects of the piece [21].
- The system should not just be a representation of a real-world system or human performer – its abstract, digital nature should be exploited [5].

2.4 Tools

The system uses physical modeling techniques with the *jit.phys* objects in the graphical programming language Max.¹ Signal

---

¹ https://cycling74.com/
² http://www.jamorigin.com/products/midi-guitar/
analysis is handled by Jam Origin’s MIDI Guitar.\(^2\) The final system is implemented as a collection of Max for Live\(^3\) devices and custom instruments within Ableton Live.\(^4\) This enabled a modular approach with efficient access to transport and mixer functions, audio effects, MIDI handling, recording and digital instruments.

### 2.5 System Description

Ten virtual bells, each consisting of a body and a clapper are modeled. Specific notes detected in the guitar signal ‘ring’ a particular bell via a virtual force (impulse) on its body. The strength of this impulse is dependent on the detected velocity of the note. This ‘ringing’ action is meant to model the action of a bell-ringer moving the body of the bell by pulling on a rope. Church bells were selected as a metaphor due to the potential of their simple workings to be recreated in software as a transparent, audio-visual system, their thematic richness for compositional exploration and their ability to produce chaotic rhythms and melodies.

When the movement of the bell causes an impact between the body and the clapper, a MIDI note is sent to samplers and synthesizers to generate the bell’s sound. To add sonic variation, MIDI notes resulting from the initial strike are separated out from the subsequent strikes caused by the bell’s motion. In addition, this action also pitch-shifts the live vocal of the human performer to create additional vocal layers according to the pitch of the sounding bell.

As shown in Figure 2, the system’s audio outputs are mixed with the live signal from the human performer’s guitar and microphone, while the visual outputs shown in Figure 1 are displayed on a screen situated next to the human performer.

A score-following mechanism, analogous to the system’s knowledge of the piece of music and based on the incoming MIDI notes, was implemented. Detection of a sectional change alters the bells’ behavior and the system outputs. A simple melody recognition system signifies the end of the piece and the bells, released from their hinges, fall to the virtual floor.

### 2.6 Development Process

#### 2.6.1 Initial Modeling

Church bells were selected as a metaphor due to the potential of their simple workings to be recreated in software as a transparent, audio-visual system, their thematic richness for compositional exploration and their ability to produce chaotic rhythms and melodies.

The basic system was implemented using jit.phys objects in Max. The ten bells of Notre Dame were used as the basis for the pitch of each bell. To enhance the transparency of the system, the virtual bells were sized and arranged according to pitch (see Figure 1). Bell collisions created MIDI notes, which were sent to samplers and synthesizers in order to create a rich, semi-electronic bell sound.

#### 2.6.2 Initial Playing

Initial playing of the system was with a MIDI keyboard. This enabled the author (a vocalist/guitarist) to trigger individual bells and listen to the outputs of the system and make refinements to it without the physical restrictions imposed by holding and playing a guitar. As well as ensuring the model behaved in a satisfactorily realistic way, this stage enabled the system to influence the establishing of the essential parameters of the composition such as tempo, note density and to some extent, mood.

#### 2.6.3 Jamming

The next stage was to jam with the system on both vocals and guitar, as if improvising with another human musician. The guitar input (analyzed by the MIDI guitar software) was used in place of the MIDI keyboard. No additional sensors were used; allowing the human performer to concentrate on developing initial picking patterns, chord changes and vocal melodies that effectively complemented the system’s output, while also feeling a strong sense of control over it.

#### 2.6.4 System Refinement

Following this, the system was further refined to complement the composed material. For example, in order to better complement the guitar part, the four highest pitched bells were linked to multiple guitar notes (Figure 3) so they would be triggered more frequently and create a more complex response.

The lower-pitched bells were mapped to individual bass notes in the guitar part so that they would better complement the chord changes and increase system transparency. To further enhance the causal relationship between the guitar and bells, the two sets of bells were separated in terms of audio outputs and visual arrangement.

---

\(^2\) http://www.jamorigin.com/products/midi-guitar/

\(^3\) https://www.ableton.com/en/live/max-for-live/

\(^4\) https://www.ableton.com/en/live/

---

**Figure 1**: The ten virtual bells

**Figure 2**: Overview of the system

**Figure 3**: MIDI note to bell mappings
In order to further develop the sonic complexity of the bells, the velocities of bell collisions were detected and linked to MIDI instrument parameters. In addition, a distinction was made between primary and secondary collisions to enable these MIDI events to be sent to separate banks of instruments, again increasing the complexity of the system’s response.

To allow the system to directly influence the human performer’s output (and therefore be truly interactive), the notes produced by the bells controlled a pitch-shifting effect (using the retune~ object) on the live vocal.

**Figure 4: Retune effect placed on live vocal channel**

### 2.6.5 Developing and Fixing the Composition

Further jamming allowed these system developments to influence the composition of the first, slower section of the piece. This included the refinement of the guitar part to better fit with the bell sounds and deciding on sounds and durations of possible lyrics as informed by the new backing vocal layers. Lyrics were then written that were inspired thematically by the metaphor and sonically by its realization.

### 2.6.6 Dynamic Behavior and Mappings

Following the fixing of the first section of the piece, the system was experimented with further to assess the possibilities of incorporating new behaviors and mappings that exploit the machine performer as a digital system. The rotational constraints on the bells hinges were removed, and impulses from guitar notes were replaced with motors that caused the bodies of the bells to spin through 360 degrees. This action suggested the use of church bells to warn of invasions during wartime, and so the outputs were mapped to synthesizers designed to resemble air-raid sirens.

In a further extension to the metaphor, a mechanism was implemented to temporarily disable the hinge holding a bell’s body when the corresponding guitar note was played above a pre-determined velocity threshold. This caused it to collide with other bells and the walls of the virtual world, as illustrated in Figure 5. These events were used to trigger additional sounds.

### 2.6.7 Score-Following

Pitch-based score-following using the incoming MIDI stream and an array of **match** objects was implemented to enable the system to detect the slower section, the faster second section and the ending of the piece. This enabled the automatic transition to the mode described above (and back again). A double-time kick drum was used to emphasize the transition to the faster section and mirror the more chaotic mappings of the system. A simple melody recognition system, triggered at the end of the piece, causes the bells to fall and produce a definite conclusion.

### 2.6.8 Recording

To further refine the piece, it became necessary to record the vocal and guitar parts along with the control messages from the system. This stage marked the point at which development of the system stopped, and afforded a more objective, holistic standpoint from which to carry out detailed composition, sound design and mixing work.

![Figure 5: Bell behavior in the 2nd section of the piece](image)

### 3. REFLECTIONS ON THE DEVELOPMENT PROCESS

#### 3.1 System-Building as Composition

Although there are significant differences from Chadabe’s notion of “interactive composing” [7] (mainly the non-real time compositional activity), the iterative activities of system building and playing with the system were key in the emergence of the composition.

In this project, the additional stages of significantly refining the system based on playing with it and the more detached activity of fixing certain aspects of the piece were also significant. Chadabe compares interactive composing as like sailing a boat, and warns of its highly challenging nature [7]. At certain stages in this project, trying to refine a system while playing with it and fixing compositional ideas felt like trying to turn an oil tanker! However, the significant reward of coming through these challenges is the creation of a system that significantly enriches the composition and can be performed easily.

#### 3.2 Musician-Machine Interactions

The three modes of real-time interaction (instrumental, ornamental and conversational) [16] were experienced during the playing with the system. During the early stages of the composition, when the system’s outputs were more influential in shaping the music, substantial performative agency was evident.

There were also stages when key developments arose from the memetic agency [5] of the system, such as the development from the church bells metaphor to one of air raid sirens.

Despite this significant agency, it has to be acknowledged that as in other co-performance systems, the human agent retains the status of dominant partner [7,23] who designs the parameters of the interaction and makes aesthetic choices that limit the freedom of the machine.

The following sections locate the human and machine agency and modes of interaction against key stages and activities in the development process described in section 2.6.

#### 3.2.1 Memetic Agency

During the early stages of the composition, the decision to use the metaphor of church bells facilitated the system’s memetic
agency in terms of setting the pitches of the bells, suggesting pitch and velocity input mappings and output timbres.

Whilst significant human control was exercised during system refinement to enable better interaction of the system with the guitar part, the system demonstrated memetic agency by suggesting the strengthening of the metaphor through reporting the strength of collisions and mapping these to audio outputs.

The development of dynamic mappings to extend the metaphor resulted from direct experimentation on the system, rather than playing music with it. This had a significant impact on the direction of the composition, as it led to the creation of new guitar and vocal parts, a switch to double time and the use of additional samples.

Finally, there were some artifacts of the system produced by some questionable programming methods (the author is not a computer scientist!) and choice of tools. For example, at the end of the piece the last clapper to fall is flashing – the result of a workaround to ensure that each clapper stays within the body of its parent bell when the hinges are activated.

3.2.2 Performative Agency: Conversational Mode
Triggering the individual bells with a keyboard allowed a conversational mode of interaction. The system demonstrated performative agency by establishing key, tempo and mood of the piece, as well as the level of input from the human performer.

This continued with the introduction of vocals and guitar. Initial ideas about performer input were developed into a guitar part and vocal melody that matched the established song parameters. The additional vocal layers generated by the system influenced the durations, melody and lyrics (words were chosen for sonic above semantic properties) of the main vocal part.

3.2.3 Performative Agency: Ornamental Mode
As described above, the system’s outputs had a strong influence over the piece during the earlier stages of composition. In the later stages, this agency decreased and both the human and machine performers’ parts became increasingly fixed.

However, a successful performance of the piece still requires the human performer to carefully listen and respond to the machine output in real-time, as natural human performance variations ensure that the machine’s output is never completely predictable.

3.2.4 Instrumental Mode
Finally, the system can be described as showing instrumental properties, as although delayed in time, playing specific notes on the guitar produces a reasonably predictable system response.

3.3 Liveness
With reference to Sanden’s theory, the Church Belles system and piece can be said to demonstrate liveness in several ways [21]. Temporal and spatial liveness are created through the projection of the virtual metaphor reinforcing the perception of separate human and machine performers, present in time and space.

Liveness of spontaneity is apparent through all audible events being triggered in real-time, a strong element of risk and the uniqueness of each performance. Corporeal liveness is evident in the link between the movement of the bells and the resultant sound. However, this can become less clear as the metaphor develops and the density of sounds increases.

Interactive liveness is demonstrated through the response of the system to the guitar input and the system’s influence over the live vocal sound. Finally, despite a high degree of audio and visual separation of the human and machine performers, virtual liveness could be said to arise from the human-machine duality inherent in the robotic, repitched vocal layers.

4. CONCLUSIONS
The Church Belles system and piece has successfully met the original aims and design criteria outlined above. Complex real-time and non-real-time musician-machine interactions have resulted in significant machine agency in the production and performance of a satisfying composition.

Although the final presentation of the piece is interactive (and an element of unpredictability ensures that no two performances will be the same), the system’s agency is primarily memetic. The iterative stages of system-building and then playing with it led to a gradual fixing of both the human and machine parts, which meant that the system’s performative agency was largely restricted to the earlier stages of the compositional process.

Finally, the system and piece demonstrate how, through the use of a dynamic, real-world metaphor, a machine performer can be represented as a separate entity with a high degree of liveness.

5. RECOMMENDATIONS
Though designed for a single composition, the system could easily be adapted for other works and for use by other composers. Within the main component of the system, a Max for Live device, up to ten bells can be activated, repitched, resized and repositioned through a simple user interface.

MIDI outputs are also highly configurable, again without needing to click on the “edit” button of the Max for Live device. However, the reuse for other pieces may necessitate the extraction of certain algorithms from the system that are intrinsic to the composition.

6. PERSPECTIVES
6.1 Computational Efficiency
Partially through the use of Max for Live, and partially due to lack of programming expertise, the system is not as computationally efficient as it could be. As discussed in 3.2.1, it features some questionable problem-solving methods, leading to unintended glitches.

Future incarnations of the piece (and indeed future works by the author) could benefit from collaboration with expert programmers and the use of games engines. However, although this might enable the musician to concentrate on more musical aspects, there is a danger that outsourcing some of its
development may dilute the machine influence on the compositional process.

6.2 Visual Aesthetics
One area that could definitely be improved by collaboration would be in the visual design of the system. Although successful as a simple schematic that does not detract from the music, some input from a visual artist might help the machine to achieve the "technological sublime", where it becomes an aesthetic object in its own right [10].

6.3 Audio Interactions and Outputs
The bells metaphor suggests further exploration of timbre through exploration of partials. For example, partial detection in the guitar signal could be used to trigger several bells simultaneously. The exact pitch produced by each bell could be dependent on the fundamental frequency, leading to additional harmonic complexity.

6.4 Spatialization
Returning to the original metaphor for further inspiration, the Notre Dame bells are positioned in different parts of the cathedral. These locations could be visually represented in three dimensions, reinforced with spatialized audio.

6.5 Dynamic Mappings
The flexibility of the system in terms of making adjustments to the bells behavior in response to score position could be further exploited in a piece with more sections. The bells’ pitches, sizes, positions, masses, friction, restitution and movement can all be modified within a piece. Exploiting further the idea of dynamic mapping affords greater compositional complexity without compromising the transparency of the system to the audience.

7. REFERENCES

8. APPENDICES
A video of a live recording of the piece with the system visuals can be found at https://vimeo.com/152642239.