ABSTRACT

This paper describes research, investigations, creative experiments and performances undertaken by the author in collaboration with practitioners in different creative and performance domains. The research focuses on the translation of expression between these domains and its implementation using technology. This paper focuses primarily on the role of notation in this process. The domains involved include music (audio and notation), movement (dance) and text (poetry). The data arising from performers’ movements are collected and investigated; consideration is given to the use of image and graphics enabling elementary algorithmically generated dance notation.

These implementations are taken to be a part of the creative process. This research is about creating and investigating stimulating experiences where connections between one domain and the other are perceivable and where this connection itself provides an aesthetic experience. They are not intended to be fixed and permanent (although may remain so for the duration of a composition). The research is about creating dynamic environments, not musical instruments or general purpose tools.

1. THREE STREAMS

1.1 Algorithmic generation of material

The practice-led research described here is the result of the concatenation over time of a number of research strands, the first of which is the algorithmic generation of material. My primary interests involve music notation but through collaborative work these have widened to include text-based material - mainly poetry - as well as the consideration of image and graphics-based work involving notations such as dance (e.g. labanotation) and the more graphical components of music notation.

It is important to note that this work does not currently attempt to use artificial intelligence, only relatively simple algorithms and physical data to generate music in ways that one might compare to traditional composition techniques.

1.2 Physical computing

The use of physical computing - physical performance in computing environments - forms a second research strand. It is necessary for the implementation of embodied expression and translation between expressive domains as well as other factors such as synchronisation in live performance and within groups. It plays an essential role in domains such as music and dance where physical effort is of significance.

1.3 Live notation

A third strand and the main focus of this paper is notation (in this case music and text) and in particular with regard to live environments. In part due to the growth of popularity of middleware such as Open Sound Control (OSC) which facilitate bespoke communications between hard and software environments, and also because of technological and in particular network-based innovations, there are increasing technologies allowing live control over a variety of notations. One of the most visible examples of these is Google Docs, but software such as INSORE [1] provides a variety of specialised notational and graphic tools, designed to be solely controllable using OSC (and therefore over networks). Related software includes MaxScore [2], the Bach Project [3] and Quintet.net [4]. While these packages each has their own advantages, they do not share INSORE’s focus on control over and flexibility in graphical presentation which is particularly important in the author’s implementation of notation synthesis for live performance.

By concentrating on the presentation and interpretation of notation, INSORE encourages freer, more intuitive methods of composition using small, ‘local’ algorithms that together generate material such as that shown in Figure 1 - material generated in response to dancers’ physical movements. These phrases are not generally pre-composed (although they could be - this is a choice made driven by aes-
thetic and practical considerations: the musicians are quite happy to encounter the music in this way). INSCORE also allows considerable control over the presentation of notation, an important feature for those composers who, like the author, find the appearance of notation reflects its expressivity (while being mindful of notation devotee Cornelius Cardew’s warning that ‘a musical notation that looks beautiful is not a beautiful notation, because it is not a function of a notation to look beautiful’ [5]).

1.3.1 Live text

Unsurprisingly, ‘liveness’ has different consequences in different domains. For those working in the domain of text the ability of Google Docs to update material synchronously for all users is literally a demonstration of the editing of material as ‘performance’. Inevitably some creative artists have used this platform as a way of interrogating particular methods of creating, viewing and performing with text [6]; others have used features of Skype and Twitter in similar ways [7].

Book publishing tends to emphasise the finished product - the messy processes of writing and editing are obscured by the impeccable published item. There have been a number of projects making use of electronic and networked resources, including novel-writing as performance [8] and as real-time performance [9], writing as performance art [10], writing as a contest against time [11] and against other authors on-line in the Penguin Books competition ‘We Tell Stories’ [12].

Of course text can also be created and manipulated generatively rather than collaboratively. This is less prevalent in text-based media (although ‘off-line’ methods such as Oulipo [13] are well known and understood). One of the first practical references to the possibility of the algorithmic generation of meaningful text was by Alan Turing [14]. In this famous test Turing replaces the question “can machines think” with “are there imaginable digital computers which would do well in the imitation game?” (The imitation game is one possible implementation of the Turing test.) While the test is for intelligence, in effect a major factor in communication is the requirement for the proper parsing of grammar through algorithms.

This apparently simple idea has been highly influential as well as controversial. In 2014 the press reported ‘the first computer ever to pass the Turing Test’ [15] - a claim quickly disputed [16]. Eugene Goostman [17] joins a long list of attempts at the algorithmic generation of meaning, stretching back through chatterbots such as ELIZA [18].

More recently there has been interest in the generation of robotic or virtual algorithmic creatures, for instance examples of real-time animation Larry the Zombie [19], or Milo from Kinect [20].

Through these examples and others it is clear that live action requires a particular aesthetic - books, films, art and music are all based on planning or improvisation. Live action/live art tends to be based on forms of guided improvisation or semi-improvisation with forms that were not previously available, so allowing hybrid creative structures involving group and real-time coordination through generative notations.

1.3.2 Live notation in music

Music, drama and dance are temporal art forms having significant improvisatory and/or interpretive components.

Over the last fifty years particular emphasis, even reverence [21], has been placed on the ‘urtext’ - most obviously in ‘classical’ musics where the score is, or has become, a fundamental element. This contrasts with many popular musics and jazz where the skilful variation or personalisation of an existing ‘standard’ is frequently considered central (witness Bob Dylan’s own increasingly inventive variations in his performances of Like a Rolling Stone). In classical musics performers have been vilified for veering too far away from the original instruction or a ‘classical’ interpretation [22]. In forms where scores are less definitive - pop, jazz and other oral, aural and more improvised forms, ‘liveness’ is not in the form of notation, but in musical signals passing between musicians. (It may be significant that so-called tribute bands - replicas of older pop acts - now exist for whom authenticity is now a main criteria.) All of these factors make the live generation of music notation a particularly hybrid form. Classically-trained instrumentalists are readily able to create dynamic and exciting performances from carefully constructed live notation - they are used to creating performance in deplorably short spaces of time from fearsome scores, after all. In this case, the live notation should not be too difficult and proper thought must be given to its format and presentation (how to judge when to ‘turn a page’ - whatever that means digitally - for instance). The author’s experience is that under these conditions musicians find performing from live scores exciting and exhilarating [23].
In the technical operation of algorithmically structuring notation it is of prime importance to achieve a satisfactory balance between the maintenance of musical style and the creation of notation straightforward and clear enough to enable the musician to give an expressive performance even when almost sight-reading. For this reason the author has made the choice to stick primarily to common practice notation. In addition, the notation has been kept as simple as possible bearing in mind the modernist style of the music. These choices have been made in order to facilitate the skills of classically-trained performers who have, through years of experience, a particular relationship with notation and they are able to transform it into dynamic, expressive performance.

Nonetheless, the live generative use of music notation has been generally less visible. While software for music notation has been developing for many years (Notator and Finale in 1988, Sibelius publicly released in 1993), there has been little apparent interest in methods of using notation both generatively and in live environments. More recently, Lilypond (e.g. [24]) has been used extensively as a platform for non-real-time generation of notation and systems such as PGWL [25] and Slippery Chicken [26] have added very sophisticated notation facilities to computer-aided-composition software. As mentioned in section 1.3 there are now a number of options available to composers working in live music notation ( [2–4, 27]), although the emphasis of both remains on computer-aided composition.

Prominent ‘historical’ examples of live notation in music include Baird [24], Wulfson [28] and Kim-Boyle [29]. The use of notation in these cases mainly consists the manipulation of image files or the generation of large quantities of material - for instance through the algorithmic coding of Lilypond files [30]. However there are some more significant uses of live generated scores [31, 32]. Volume 29:1 of Contemporary Music Review (2010) is given over entirely to a review of live notation.

Unsurprisingly in a comparatively new field there are significant issues yet to be dealt with in the practical implementation of live notation. These include bridging the technical and aesthetic divide between notation and signals [31], general complications with synchronisation and timing, practical difficulties such as when to ‘page turn’, how to achieve the correct balance between reading and improvisation as well as inherent issues such as sight-reading and how difficult-to-play notation can become before it requires practice. As Lukas Foss commented on, "the precise notation which results in imprecise performance" and that "to learn to play the disorderly in orderly fashion is to multiply rehearsal time by one hundred" [33].

1.3.3 Live notation in dance and graphics

Prominent extant forms of dance/movement notation include Labanotation, or Kinetographie Laban by Rudolf von Laban [34], Benesh Movement Notation (graphical representation of human bodily movements), Eshkol-Wachman Movement Notation (graphical representation of bodily movements of other species in addition to humans, and indeed any kind of movement (e.g. aircraft aerobatics)) as well as others. These forms are primarily graphical reflecting their main focus on movement rather than textual or symbolic meaning.

While some forms of music notation have had a long and varied history, dance notation has not been so prominent. One of the reasons for this lies in the different functions that exist for dance notation. It is usually considered as a way of storing and passing on existing dances rather than as a way of expressing oneself, making the adoption or even exploration of dance or movement notation more difficult. It is rarely used in the communication of new dance work, and in spite of Albrecht Knust’s suggestion that in Labanotation “the symbols must speak directly to the eyes of the reader so that he can perform the movements without, or at least without too much, reflection” [35], there are questions as to how easily and quickly it can be read and digested. Text and music notations are generally so well understood by performers that this is not a problem (although it usually requires some time to ‘digest’ them (see section 5)). Some musics have tests for sight-reading ability, implying that financial considerations are very likely to reduce the capacity for detailed rehearsal!

A further difference is that dance notation is generally considered such a specialised field that professional notators need to be employed, limiting its take-up in live work.

Finally, a problem specifically associated with the live use of this notation is how it can be communicated to the dancers. Most commonly this is via a data projector, but this limits the dancer’s movements significantly.

Recent developments linking live notation and dance have included a variety of instances of ‘hacking choreography’ and ‘live coding’ involving dance and other forms of embodied expression. While predominantly extensions of the physical computing methods mentioned above, the use of live coding as a form of notation has been imaginatively investigated by Alex McLean and Kate Sicchio in [36–38] and demonstrated in 2013 [39].

While there are some practical problems with these systems - mainly around communicating the notation to the dancer, McLean’s version of Texture, demonstrated in [39] is both visually striking and expressive. It does however, become increasingly complex as the dance progresses, making interpretation a particularly vital part of the interaction.
While the present condition of dance notation can appear to be quite frustrating, particularly in its lack of standardisation, the field is open for further developments in notation systems.

2. CROSS-DOMAIN EXPRESSION

These three research streams together allow for the practice-led investigation of cross-domain expression. Cross-domain ways of thinking are so natural to us that it is difficult to imagine expression without them. Performed music is itself a cross-domain activity utilising both physical and mental dexterity. (Arguably the use of mixed metaphors (such as my own use of the phrase ‘mental dexterity’ in the previous sentence) is another example, as are metaphors and analogies themselves.)

Writing about music often requires the use of metaphors and particularly when we are seeking to analyse or describe less embodied musical forms, such as acousmatic music, we are even more reliant on other domains such as language and image [40].

Most expressive domains themselves comprise of a number of linked sub-domains. A lot of music, for instance, can be described as expression through pattern enabled by physical effort. This research leans heavily on these interdependencies, seeking to maximise expression and interaction through the exploitation of musicians’ learned performance skills articulated through common practice notations.

3. SEMAPHORE

Semaphore is a collaborative music-dance-text piece composed using research which seeks to translate between expressive domains using technology. An expressive domain is a form of artistic expression such as music, dance or text. Uniquely, information is taken from one domain and translated into another in real-time so allowing simultaneous performance. The music, environment and programming is by the author, choreography is by Jane Turner and text is by the novelist and poet Phil Terry. The music is performed live from code in the SuperCollider audio programming environment [41, 42], a combination of prepared functions and structures and including some methods related to live coding.

3.1 A cross-domain sequence explained

Semaphore is composed of patterns of interactive cross-domain scenes and sequences. The following is an example of a single synchronous sequence:

A dancer’s physical movement triggers and modulates the computer generation of a text phrase, which is displayed and performed. This performance is recorded and the recording is analysed spectrally. The results of the analysis then trigger and modulate a musical phrase presented as music notation which is then played by an instrumentalist. A dancer responds to the performed phrase with a physical gesture.

This set of actions might take place over a period from a few milliseconds to one or two seconds, or over an even more extended period of time. We find that the only significant latency occurs as performers consciously respond to newly displayed notations.

Alongside its creative potential, this research enables people working in one domain to generate material in another. These people might be expert performers in another domain or members of the public with no particular expertise.

There are many examples of movement-based interfaces for music, but this work is unique in its facilitation of translations from one domain into the notation of another: music, text, dance or graphics. The use of notation allows us to preserve performance interpretation that many audience members find so fundamental in live art.

Of course, the creative problem of how to create meaningful expression from these technical procedures remains as crucial as ever.

4. TECHNICAL PROCEDURES

In the following sections ways in which the parts of the sequence described above were implemented technically outlined in more detail.

4.1 A dancer’s physical gesture...

The ubiquitous Microsoft Kinect (Xbox 360 version) is used to capture a dancer’s physical movements. The software used for programming the audio environment, SuperCollider, is also used to perform some rudimentary movement detection. Gesture recognition is not central to this research and the software does not seek to make precise distinctions between different gestures but it is used to detect the speed and range of the movements of certain body parts. Effective though the Kinect is, the Loie Fuller Ap-...
in collaboration with the team by the writer and poet Phil Terry especially for this performance. One of the key questions was how to achieve an expressive balance between sound and meaning in the text. Terry is well-versed in Oulipo techniques [13] and was aware of many possible technical textual procedures and their results - we wanted something focused and related to the Semaphore concept. Eventually, we decided on material that fell in between sound and semantics, and which also enabled some algorithmic manipulation. (Apparently by chance - or euphony - the word ‘semantic’ appears in the poem, linked sonically to ‘semaphore’.)

Semaphore or some are for just as elsewhere
some are against
Some fear to offer or seem to fear
Afar a fir so that through the undergrowth and
across the map
A flare or a car

Soars to see the same semantic dance
Oars soar with ease or seem to soar
The same flares through the firs
Seem to spore

Ears arms too as a sheer harm
Verse as shame same sheep
Sham spheres or spare harems reap hope
Marsh shears or fennel ash

When we discovered that the original poem was too short, Terry expanded it, using a pantoum structure derived from the Malay pantum verse form which repeats lines in a pattern, effectively doubling the original length:

A B C D
B E D F
E G F H
G I/A/C H J/A/C

This produces verses with a gentle, somewhat zen-like quality, emphasising the rather surreal nature of the original verse:

Some fear to offer or seem to fear
Soars to see the same semantic dance
A flare or a car
Oars soar with ease or seem to soar

Soars to see the same semantic dance
The same flares through the firs
Oars soar with ease or seem to soar
Seem to spore

Ears arms too as a sheer harm
Verse as shame same sheep

While the final part of Semaphore revolves around a pre-written poem (see section 4.3), an introductory, more abstract section (figure 3) originally involved direct interaction between dancers and text. As an example we arranged a passage where if the movements of one of the dancers was faster/higher than a given threshold, a trigger is sent to an algorithm which then chooses one from a group of selected words from the poems (such as flashing, shear, roar, billows, swelling, etc.). Although the metaphors chosen here seem rather trite or simplistic, the scenario proved expressive, successful and full of potential.

4.3 ...the recording is analysed...

For the last part of Semaphore, we recorded Terry reading the poem. As we needed to mix between dry and wet audio streams we used a recording, although the use of a live voice (at least in part) reading live generated text is a important goal.

The software analyses the frequency and amplitude components of the vocal. The base frequency generates a series of sustained chords accompanying the voice gently in the
Figure 4. Conversion process from data to audio and notation formats

background. If the frequency pushes over a certain threshold, a small melisma is triggered. Similarly, if an amplitude threshold is broken, a sharper, more dissonant chord is generated.

4.4 ...a musical phrase presented as notation...

At specific times during this episode - after about every thirty seconds or so - a snapshot is taken of the voice’s frequency. This frequency is used in the generation of the notation of sustained notes for the clarinet and ‘cello (see screenshot in Figure 5). These are arranged to create an effect in imitation of the sound of the bell of a navigational ocean buoy. In all these cases INSCORE is used to present the notation. INSCORE is controlled through OSC messages, allowing a tight integration between the language used for algorithmic control (in this case SCLang, but it could be any other OSC compatible environment) and processes synthesising the notations (see Figure 4).

4.5 ...performed by an instrumentalist and interpreted by dancers

As these notes are performed instrumentally, they are interpreted by the dancers as a port of the overall choreography. In turn, these movements may contribute further to the process of text and music notation generation. In future, we plan to use this ‘audio feedback’ to modulate the generation of dance notation (section 1.3.3).

5. LATENCY

The subject of latency frequently arises during discussions concerning performance using these technologies. Latency is defined as the time taken from the moment one event happens - in this case, the movement of a dancer - to the moment that the effect of that event is perceived - in this case, the generation of the notation and its subsequent performance [44]. The origin of the problem of latency in digital systems lies in the field of audio production and reproduction - it is the (inevitable) result of digital systems where data must be read from memory to be converted into sound. The larger quantity of data that can be read, the more efficient and to that extent the faster the system, but the higher the potential latency. Designers of digital audio instruments must find a balance between these two incompatible goals. There are, of course, examples of non-digital or mechanical latency, the time that an organ pipe (especially very large lower pitched ones) take to activate following the mechanical pressing of the key in common with many other larger acoustic instruments (double bassoon, baritone saxophone, etc.) for instance.

5.1 Causes of Latency in Semaphore

When using the Kinect, apart from the unit itself, once data is transferred to SuperCollider there are a number of additional factors that can cause latency. Most algorithmic processes dealing with symbolic musical structures (such as notation and musical pitch) will involve rather minimal processing and so will not usually cause any delay. However, the production of the notation itself can have a significant effect.

Mirroring the description of digital latency above, synchronisation with physical events requires a ‘sampling’ of those events in order to process them. Any system then has to balance the accuracy of this sampling against other system requirements. When including physical movements, especially those created through skilled dancers, we usually wish to identify general gestures rather than small movements - the upward rapid sweep of an arm, for instance. In order to achieve this we need to average the incoming data so smoothing out any sudden extraneous movements. (Of course, in some circumstances this is not wanted, in which cases the sampling windows must be kept small.)

These movements must then be mapped to musical gestures in one way or another. The author has chosen to develop these mappings [45–47] as an integral part of the creative process. They may be very straightforward one-to-
one mappings [45, 46] - for instance an upwardly moving arm might produce an upwardly proceeding arpeggio or scale - or it may be used as a form of gesture - a fast movement may produce a fast moving string of notes (see notes 2-5 in Figure 1 above). Equally the mapping may include some aspects of real-world behaviours and gestures [48].

In some cases it is not possible to conclude a musical phrase without synchronous information, again meaning that some form of latency is inevitable.

Finally, the involvement of humans and human perception and notation is itself probably the greatest cause of latency. Rehearsals with live notation suggest that ideally performers need a second or so from the moment that the new notation is displayed to properly digest and respond to it.

5.2 Effects of Latency

Stimulating creative results seem to arise from these developmental, even compositional choices, sometimes emphasising a direct, easily perceptible relationship between movement and result, sometimes confounding expectations with a melismatic flurry as if from nowhere.

One of the difficulties some have with high levels of latency is that there is perceived to be a lack of control, even a lack of feeling of cause and effect. This implies that our main aim should be the creation of musical instruments in the best traditions of the New Interfaces for Musical Expression conference [49]. However, the design of musical instruments is not the main focus in this research. One aim in Semaphore is to investigate whether expert expressive movement can find a mapped reflection in another domain, in this case music or text. Latency might be a feature of the systems, but is not an issue for the team. If precisely timed responses are required, solutions are easily available, such as strict pre-planning of rhythm, movement and display or even the simple playback of recordings.

6. AUDIENCE RESPONSE

6.1 Universities’ Week

Universities’ Week provided a particularly successful occasion for about 60 members of the public of all ages to interact with our system voluntarily. Although interactions produced somewhat modernist music without clear melody or rhythm and although it is likely that only a relatively few of the participants understood music notation it was clear that most enjoyed the experience immensely. Children in particular seemed to relax and expressed themselves enthusiastically and with none of the self-consciousness so typical of their parents. A video recording of these interactions is available - please contact the author for access (see Figure 6 for an example screenshot).

6.2 Rehearsal and acquaintance with the system

Feedback on all aspects of the composition and the notation system was gathered from the participants throughout the rehearsal process. This included two early rehearsals during which the author worked with one student dancer to properly ascertain basic functionality such as sensor ranges and sensitivities. While the Kinect can be quite sensitive to some movements it is also the case that its basic design is to recognise simple bodily movements usually associated with sports and gaming rather than the sometimes delicate and gentle movements used in contemporary dance. These factors were also linked to allowances made for latency and reliability (see section 5). In Semaphore there are relatively few requirements for absolute and precise temporal coordination, although we are optimistic that more precise synchronisation can be achieved as the systems develop.

Performers were encouraged to provide informal feedback throughout the rehearsal process and, as has happened in the past, it was soon apparent that the main problems emerged not from the generated music but rather how it was displayed.

A quick comparison of paper sizes and areas (table 1) shows that the screen area of a 15” MacBook Pro is quite small - resolution is rather irrelevant as quite a large size
of notation is needed. Traditional music paper sizes are far from standardised, but tend to be quite significantly larger. The laptop’s screen also only allows for the viewing of one ‘page’ at a time and this small screen is in landscape mode rather than portrait. All these factors mean that it is a very different experience reading from a laptop’s screen rather than from pieces of paper.

Another problem relating to screen size and presentation is when ‘page turns’ should occur and in this new environment exactly what a page turn is. In paper parts page turns, particularly those parts where frequent or near constant playing is demanded, are planned carefully, maximising the time available to turn the page at the most convenient moment. This also means that when a musician turns the page they can consciously ‘discard’ previous information. Semaphore attempts a variety of experimental solutions, none of which are optimal as yet.

At the moment it is clear that the use of live notation requires compromise in how it is implemented and used. For some composers these compromises may simply be too radical to consider at present.

Jonathan Eacott [50] suggests that there is a requirement in live notation for ‘a metronome or cursor to keep musicians in sync’ and that there ‘must be a way of continually scrolling the music so that musicians can look ahead’ - these features would certainly be very useful. However, they are not essential, depending on the nature of the material presented. If the music appears note by note as it is being created this has the advantage that it can give a fairly clear indication of the ‘tempo’ at which it should be played, and any further synchronisation can be achieved between instrumentalists as usual: paper parts do have cursors or metronomes.

Apart from these issues, all involved with Semaphore and earlier pieces such as Calder’s Violin have been very positive about their experience with. Although some have displayed confusion and anxiety on first acquaintance, after some rehearsal and after realising that they are not required or expected to play every note with perfect accuracy, they relax and even enjoy the experience [23].

7. CONCLUSIONS

All who have been involved in Semaphore have been gratified by the response received from audiences and workshop visitors. The audience were offered the chance of completing a general questionnaire; fourteen were completed. These were uniformly positive; a number also contained free text comments. Below are included a selection of these, included not in a spirit of self-congratulation, but in order to demonstrate the connection felt between audience, the dancers’ physical movements and the resulting music, both audio and notation:

- “I really enjoyed the performance... it was interesting to watch the dancers ‘create’ the music.”
- “I came because of a fondness for dance but ... there is so much to take in here that it was useful to have to have two performances of the piece... Another couple of renditions would have permitted me to take in fully the choreography, the score, the text and the interaction of all the elements.”
- “Thanks, it was beautiful”
- “Very interesting, would attend another similar event”
- “Really engaging and interesting... [the] performance was captivating”
- “It was great, and I wish more events had a discussion and then second performance format, that worked well”
- “Brilliant!”

Those who took part during the Universities’ week also clearly demonstrated that people find generating music in this way very enjoyable and rewarding. There would also appear to be a deep link between the domains of physical movement and music. Semaphore shows that it is also possible to create and manipulate translations between music, movement and text and that both performers and audience find this expressive and stimulating. We very much hope to continue to develop these systems to enable expression and experimentation between domains. There are many possibilities that we have not even yet begun to explore.

Acknowledgments

The author would like to thank Anglia Ruskin University, Arts Council England and Turning Worlds Dance for their support in this project. In addition I would particularly thank the Jane Turner, Phil Terry, Kate Romano, Cheryl Frances-Hoad, Ann Pidcock, Gwen Jones and Ciara Atkinson for their support.

8. REFERENCES


