Brain-computer music interface for generative music

E R Miranda

1Interdisciplinary Centre for Computer Music Research, University of Plymouth, Drake Circus, Plymouth, UK

eduardo.miranda@plymouth.ac.uk
cmr.soc.plymouth.ac.uk

ABSTRACT

This paper introduces a brain-computer interface (BCI) system that uses electroencephalogram (EEG) information to steer generative rules in order to compose and perform music. It starts by noting the various attempts at the design of BCI systems, including systems for music. Then it presents a short technical introduction to EEG sensing and analysis. Next, it introduces the generative music component of the system, which employs an Artificial Intelligence technique for the computer-replication of musical styles. The system constantly monitors the EEG of the subject and activates generative rules associated with the activity of different frequency bands of the spectrum of the EEG signal. The system also measures the complexity of the EEG signal in order to modulate the tempo (beat) and dynamics (loudness) of the performance.

1. INTRODUCTION

Research into Brain-Computer Interface (BCI) for music is an interesting arena for the development of new possibilities in recreational and therapeutic devices for people with physical and neurological disabilities; we shall refer to these systems as Brain-Computer Music Interfaces (BCMI). There are various music-making devices available for those with disabilities, and even though these devices have proved to work very effectively, they often do not allow as much control for those with severe physical disabilities. At present, access music tutors use gesture devices and adapted accessible technology to make this possible, which achieve excellent results; although for people with severe physical disabilities, having complete control of the environment created for them by the facilitator can sometimes prove difficult. This paper introduces a BCMI that uses the EEG steer generative musical rules in order to compose and perform music on a MIDI-controlled mechanical acoustic piano (Figure 1a).

Human brainwaves were first measured in 1924 by Hans Berger, who termed these measured brain electrical signals the electroencephalogram (literally “brain electricity writing”). Berger first published his brainwave results in 1929 (Berger 1929). Today, the EEG has become one of the most useful tools in the diagnosis of epilepsy and other neurological disorders. Further, the fact that a machine can read signals from the brain has sparked the imaginations of scientists, artists and other enthusiasts, and EEG has made its way into applications other than clinical. In the early 1970s, Jacques Vidal, a researcher at the University of California Los Angeles, did the first tentative work towards a brain-computer interface (BCI) system. The results of this work were published in 1973 (Vidal 1973). Despite a number of isolated initiatives at building BCI systems, research in this area did not take off until the early 1990s, probably due to technological limitations. In 1990, Jonathan Wolpaw and others at the Wadsworth Center at State University of New York Albany and Rensselaer Polytechnic Institute developed a system to allow some control of a computer cursor by individuals with severe motor deficits. Users were trained to use aspects of their EEG to move a cursor on a computer screen (Wolpaw et al. 1991). In 1998, Christoph Guger and Gert Pfurtscheller (Guger and Pfurtscheller 1998) presented a paper at the 9th European Congress of Clinical Neurophysiology reporting impressive advances in BCI research: an EEG-based system to control a prosthetic hand. Many attempts followed with various degrees of success. For recent reports on BCI research please refer to the special issue of IEEE Transactions on Biomedical Engineering published in June 2004 (Vol. 51). To date, most efforts of BCI research have been aimed at developing assistive technology to help people communicate via computer systems and/or control mechanical tools, such as a wheelchair or a prosthetic organ. Comparatively little has been done to address the use of BCI technology for musical applications; such applications could
undoubtedly enhance the quality of life for people with disabilities in many ways, be they recreational, therapeutic or professional.

As early as 1934, a paper in the journal *Brain* had reported a method to listen to the EEG (Adrian and Mathews 1934), but it is generally accepted that it was composer Alvin Lucier, in 1965, who composed the first musical piece using EEG: *Music for Solo Performer*, a piece for percussion instruments played by the resonance of the performer’s EEG. Lucier placed electrodes on his own scalp, amplified the signals, and relayed them through loudspeakers that were directly coupled to percussion instruments, including large gongs, cymbals, tympani, metal ashcans, cardboard boxes, bass and snare drums (Lucier 1976). The low frequency vibrations emitted by the loudspeakers set the surfaces and membranes of the percussion instruments into vibration. Later in the 1960s, Richard Teitelbaum used various biological signals including the EEG and ECG (electrocardiogram) to control electronic synthesisers. In the early 1970s David Rosenboom (1975) began systematic research into the potential of EEG to generate art works, including music. Drawing on concepts from Cybernetics, he developed EEG-based musical interfaces associated with a number of compositional and performance environments that used the state-of-the-art EEG technology of the day. In particular, Rosenboom explored the hypothesis that it might be possible to detect certain aspects of our musical experience in the EEG signal. For example, in a paper that appeared in 1990 in *Computer Music Journal*, he introduced a musical system whose parameters were driven by EEG components believed to be associated with shifts of the performer’s selective attention (Rosenboom 1990). Thirteen years later, this author and colleagues published a new paper in *Computer Music Journal* reporting experiments and techniques to enhance the EEG signal and train the computer to identify EEG patterns associated with different cognitive musical tasks (Miranda et al. 2003).

2. THE ELECTROENCEPHALOGRAM

The EEG is measured as the voltage difference between two or more electrodes on the surface of the scalp one of which is taken as a reference. The EEG expresses the overall activity of millions of neurons in the brain in terms of charge movement, but the electrodes can detect this only in the most superficial regions of the cerebral cortex. There are basically two conventions for positioning the electrodes on the scalp: the 10-20 Electrode Placement System (as recommended by the International Federation of Societies for EEG and Clinical Neurophysiology), and the Geodesic Sensor Net (developed by a firm called Electric Geodesics, Inc.). The former is more popular and is the convention adopted for the system described in this paper: it uses electrodes placed at positions that are measured at 10% and 20% of the head circumference (Figure 1b). In this case, the terminology for referring to the position of the electrodes uses a key letter that indicates a region on the scalp and a number: F = frontal, Fp = frontopolar, C = central, T = temporal, P = parietal, O = occipital and A = auricular (the ear lobe; not shown in Figure 1b). Odd numbers are for electrodes on the left side of the head and even numbers are for those on the right side. The set of electrodes being recorded at one time is called a montage. Montages normally fall into one of two categories: referential or bipolar. Referential means that the reference for each electrode is in common with other electrodes; for example, each electrode may be referenced to an electrode placed on the earlobe. An average reference means that each channel is compared to the average potential of every electrode. Bipolar means that each channel is composed of two neighbouring electrodes; for example, channel 1 could be composed of Fp1-F3, where Fp1 is the active electrode and F3 is the reference, then channel 2 could be composed of Fp2-F4, where Fp2 is the active electrode and F4 is the reference; and so forth.

It takes many thousands of underlying neurons, activated together, to generate EEG signals that can be detected on the scalp. The amplitude of the EEG signal strongly depends on how synchronous is the activity of the underlying neurons. The EEG is a difficult signal to handle because it is filtered by the meninges (the membranes that separate the cortex from the skull), the skull and the scalp before it reaches the electrodes. Furthermore, the signals arriving at the electrodes are sums of signals arising from many possible sources, including artifacts like the heartbeat and eye blinks. Although medical specialists can diagnose brain malfunctioning from raw EEG plots, this signal needs to be further scrutinized with signal processing and analysis techniques in order to be of any use for a BCI system. There are a number of approaches to EEG analysis, such as power spectrum, spectral centroid, Hjorth, event-related potential (ERP), principal component analysis (PCI) and correlation, to cite but a few. Although powerful mathematical tools for analysing the EEG already exist, we still lack a good understanding of their analytical semantics in relation to musical cognition. However, continual progress in the field of Neuroscience of Music (Peretz and Zatorre 2003) is substantially improving this scenario. Once these issues are better understood we will be able to recognise patterns of cognitive activity in the brainwaves and activate appropriate musical algorithms associated with such patterns. Preliminary work in this regard has been reported in (Miranda et al. 2003;
Miranda et al. 2004). Brief non-mathematical introductions to EEG power spectrum and Hjorth analyses are given below due to their relevance to the system presented in this paper.

**Figure 1.** (a) The BCMI system in action. Note the keys of the piano being played by the system; no hands or any kind of physical movement is needed. (b) The 10-20 electrode placement system.

### Table 1: Electroencephalographic rhythms.

<table>
<thead>
<tr>
<th>Rhythm</th>
<th>Bandwidth</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>Delta</td>
<td>Lower than 4 Hz</td>
<td>Sleep (non-dreaming)</td>
</tr>
<tr>
<td>Theta</td>
<td>Between 4 Hz and 7 Hz</td>
<td>Drowsiness</td>
</tr>
<tr>
<td>Alpha</td>
<td>Between 8 Hz – 13 Hz</td>
<td>Relaxed; aware but with eyes closed</td>
</tr>
<tr>
<td>Beta</td>
<td>Higher than 13 Hz</td>
<td>Awake, alertness, intense mental activity</td>
</tr>
</tbody>
</table>

Power spectrum analysis is derived from techniques of Fourier analysis, such as the Discrete Fourier Transform (DFT). In short, DFT analysis breaks the EEG signal into different frequency bands and reveals the distribution of power between them. This is useful because the distribution of power in the spectrum of the EEG reflects distinct “mental states”. In general, a signal characterised by low-frequency components (e.g., lower than 8 Hz) may be associated with a state of drowsiness, whereas a signal characterised by high-frequency components (higher than 14 Hz, up to 30 Hz or so) may be associated with a state of alertness. When the cortex is most actively engaged in processing information, whether generated by sensory input or by some internal process, the activity level of cortical neurons is relatively high, but also relatively unsynchronized. By contrast, when the cortex is less actively engaged in processing information, the activity level of the cortical neurons is relatively low, but also relatively synchronized. The higher the synchronization of the neurons, the higher the overall amplitude of the EEG (Giannitrapani 1985). There are four recognised frequency bands of EEG activity, also referred to as EEG rhythms, associated with different mental states in healthy adult brains, as shown in Table 1 (Misulis 1997).

Hjorth (1970) introduced an interesting method for clinical EEG analysis, which measures three attributes of the signal: its activity, mobility and complexity. Essentially, Hjorth analysis is a time-based amplitude analysis. This method is interesting because it represents each time step (or window) using only these three attributes and this is done without conventional frequency domain description. The signal is measured for successive epochs (or windows) of one to several seconds. Two of the attributes are obtained from the first and second time derivatives of the amplitude fluctuations in the signal. The first derivative is the rate of change of the signal’s amplitude. At peaks and troughs the first derivative is zero. At other points it will be positive or negative depending on whether the amplitude increases or decreases with time. The steeper the slope of the wave, the greater will be the amplitude of the first derivative. The second derivative is determined by taking the first derivative of the first derivative of the signal. Peaks and troughs in the first
derivative, which correspond to points of greatest slope in the original signal, result in zero amplitude in the second derivative, and so forth. Activity is the variance of the amplitude fluctuations in the epoch. Mobility is calculated by taking the square root of the variance of the first derivative divided by the variance of the primary signal. Complexity is the ratio of the mobility of the first derivative of the signal to the mobility of the signal itself. There is no clear agreement as to what these measurements mean in terms of cognition. It is common sense to assume that the longer a subject remains focused on a specific mental state, the more stable the signal is, and therefore the variance of the amplitude fluctuation is lower.

3. THE GENERATIVE MUSIC ENGINE

Our system features a machine-learning algorithm that learns generative music rules from given examples (i.e., musical scores in machine-readable format). Due to limited availability of space, this paper focuses only on the functioning of the generative component of the system.

The EEG signal can influence in a well-defined way the mixture of different style-elements found in the different musical examples given to train the system. It can generate music that contains, for example, more Schumann-like elements when the spectrum of the subject’s EEG is characterised by alpha rhythms or more Beethoven-like elements when the spectrum of the EEG is characterised by beta rhythms; these associations are arbitrary.

| Table 2: An excerpt from a database of musical elements where: CO = composer (SCHU = Schumann), P-CLASS = pitch class, P = pitch, PCL = pitch-class leading, PL = pitch leading and TYPE = type. |
|-------------------|-------------------|-------------------|-------------------|-------------------|
| ID                | SCHU-1-1-CAD      | ID                | SCHU-1-1-MEA-6    |
| CO                | SCHU              | CO                | SCHU              |
| P-CLASS           | ((0 2 7) (0 2 4 5 7 11)) | P-CLASS           | ((5 9) (0 5 7 9)) |
| P                 | 74                | P                 | 81                |
| PCL               | ((0 4 9) (0 2 4 5 7 11)) | PCL               | ((0 2 7) (0 2 4 5 7 11)) |
| PL                | 76                | PL                | 74                |
| TYPE              | CAD               | TYPE              | BAR               |
| ID                | SCHU-1-1-MEA-1    | ID                | SCHU-1-1-MEA-4    |
| CO                | SCHU              | CO                | SCHU              |
| P-CLASS           | ((0 4) (0 3 4 6 7 9)) | P-CLASS           | ((0 4) (0 3 4 6 7 9)) |
| P                 | 76                | P                 | 83                |
| PCL               | ((2 7 11) (2 5 7 9 11)) | PCL               | ((0 4) (0 4 7)) |
| PL                | 83                | PL                | 76                |
| TYPE              | INC               | TYPE              | BAR               |
| ID                | SCHU-1-1-MEA-3    | ID                | SCHU-1-1-MEA-2    |
| CO                | SCHU              | CO                | SCHU              |
| P-CLASS           | ((0 4) (0 3 4 6 7 9)) | P-CLASS           | ((2 7 11) (2 5 7 9 11)) |
| P                 | 76                | P                 | 83                |
| PCL               | ((2 7 11) (2 5 7 9 11)) | PCL               | ((0 4) (0 3 4 6 7 9)) |
| PL                | 83                | PL                | 76                |
| TYPE              | BAR               | TYPE              | BAR               |

Example-based musical-generation systems are often based on formalisms such as transition networks or Markov Chains to re-create the transition-logic of what-follows-what, at the level of notes and at the level of similar “vertical slices” of music. The act of recombining the building blocks of music material together with some typical patterns and structural methods has proved to have great musical potential (Cope 2001). Self-learning predictors of musical elements based on previous musical elements can be used at any level or for any type of musical element such as: musical note, chord, bar, phrase, section, and so on. The current version of the system uses a simple statistical predictor at the level of short vertical slices of music such as a bar or half-bar, where the predictive characteristics are determined by the chord (harmonic set of pitches, or pitch-class) and by the first melodic note following the melodic notes in those vertical slices of music. The system uses a method for generating short musical phrases with a beginning and an end that allows for real-time steering with information extracted from the EEG. The system generates music by defining top-level structures of sequences and methods of generating similarity- or contrast-relationships between elements. Consider the following example (LISP-like notation):

S -> (INC BAR BAR BAR BAR BAR HALF-CADENCE 8BAR-COPY)

From this top-level the system generates rules for selecting a valid musical building block for each symbol, including rules for incorporating the EEG information in all decisions. For example:
This defines a framework to generate a valid sequence with a beginning and an end, including real-time EEG control through the variable EEG-SET-COMPOSER. The generative engine will find a musical element for each of the constraint-sets that are generated above from INC and BAR, by applying the list of constraints in left-to-right order to the set of all musical elements until there are no constraints left, or there is only one musical element left. This means that some of the given constraints might not be applied.

The database of all musical elements contains information from different composers (or musical styles), with elements tagged by their musical function such as “measure 1” for the start of a phrase, “cadence” for the end, “composer” for the composer or musical style, and the special tags “pitch” and “pitch-class” that are both used for correct melodic and harmonic progression or direction. As an example, Table 2 lists excerpts from a database showing the main attributes that are used to recombine musical elements.

P-CLASS (for pitch-class) is a list of two elements. The first is the list of start-notes, transposed to the range of 0-11. The second is the list of all notes in this element (also transposed to 0-11). P is the pitch of the first (and highest) melodic note in this element. By matching this with the melodic note that the previous element was leading up to, we can generate a melodic flow that adheres in some way to the logic of “where the music wants to go”. The PCL (for pitch-class leading) elements contain the same information about the original next bar; this is used to find a possible next bar in the recombination process. Then there are the INC, BAR, and CAD elements. These are used for establishing whether those elements can be used for phrase-starts (incipient), or cadence. Simply by combining the musical elements with the constraint-based selection process that follows from the terminals of the phrase-structure rewrite-rules, we end up with a generative method that can take into account the EEG information. This generates musical sequences with domino game-like building block connectivity:

( (EQUAL 'MEASURE-1) (EQUAL 'COMPOSER EEG-SET-COMPOSER) )

Assuming that there are also musical elements available from composers other than SCHU, the first constraint will limit the options to all incipient measures from all musical elements from all composers. The second constraint will then limit the options according to the current EEG analysis to the composer that is associated with the current EEG activity, as follows:

( (CLOSE 'PITCH 'PREV-PITCH-LEADING) (CLOSE 'PITCH-CLASS 'PREV-PITCH-CLASS-LEADING) (EQUAL 'COMPOSER EEG-SET-COMPOSER) )

In the given phrase structure, the rule that follows from BAR then defines the constraints put upon a valid continuation of the music. These constraints will limit the available options one by one, and will order them.

Figure 2. An example of a generated mixture of Robert Schumann and Ludwig van Beethoven.
according to the defined rule preferences. The close constraint will order the available options according to their closeness to the stored value. For example, after choosing (refer to Table 2): (SCHU-1-1-MEA-1 CO SCHU P-CLASS ((0 4) (0 3 4 6 7 9)) P 76 PCL ((2 7 11) (2 5 7 9 11)) PL 83 TYPE INC) as the beginning, PREV-PITCH-LEADING will have stored 83, and PREV-PITCH-CLASS-LEADING will have stored ((2 11) (2 5 7 9 11)). This will result in measures 2 SCHU-1-1-MEA-2 and 4 SCHU-1-1-MEA-4 being ranked highest according to both pitch and pitch-class, and measure 6 and the cadence SCHU-1-1-CAD close according to pitch-class, while measure 6 SCHU-1-1-MEA-6 is also quite close according to pitch. This weighted choice will give a degree of freedom in the decision that is needed to generate pieces with an element of surprise. The music will not get stuck in repetitive loops when no perfect match is available, but it will find the closest possible continuation. The system can still find a close match this way if the third constraint eliminates all the obvious choices that are available; e.g., because a jump to the musical elements of another composer or style is requested, which might not use the same pitch-classes and pitches. Figure 2 shows an example output with elements from the musical styles of Robert Schumann and Ludwig van Beethoven. In this example, the EEG jumped back and forth from bar to bar between the two styles.

4. THE BRAIN-COMPUTER MUSIC INTERFACE SYSTEM

Our BCMI system falls into the category of BCI computer-oriented systems (Miranda et al. 2003). These systems rely on the capacity of the users to learn to control specific aspects of their EEG, affording them the ability to exert some control over events in their environments. Examples have been shown where subjects learn how to steer their EEG to select letters for writing words on the computer screen (Birbaumer et al. 1999). However, the motivation for our research departs from a slightly different angle to other BCI systems. We aimed for a system that would make music by “guessing” the meaning of the EEG of the subject rather than a system for explicit control of music by the subject. We acknowledge unreservedly that the notion of “guessing the meaning of the EEG” might be rather naïve. Nevertheless we took the risk of suggesting this notion because it is a plausible notion: it is based on the assumption that neurophysiologic information can be associated with specific mental activities (Anderson and Sijercic 1996; Petsche and Etlinger 1998). Continual progress in the field of Neuroscience of Music is increasingly substantiating this assumption (Peretz and Zatorre 2003).

The BCI research community understands that a BCI system is a system that allows for the control of a machine by explicitly imagining the task(s) in question; e.g., control a robotic arm by thinking explicitly about moving an arm. This is a very difficult problem. The system presented in this paper does not address this type of explicit control. Moreover, we are not interested in a system that plays a melody by imagining the melody itself. Rather, we are interested in furnishing our systems with Artificial Intelligence in order to allow them to make their own interpretation of the meaning of the EEG patterns. This is a fundamentally different approach to BCI because such machine-interpretations may not always be accurate or realistic. However, this is exactly the type of man-machine interaction that we are interested in exploring. The system is programmed to look for information in the EEG signal and match the findings with assigned generative musical processes corresponding to different musical styles. As mentioned in the previous section, these assignments are arbitrary.

The EEG is sensed with seven pairs of gold EEG electrodes on the scalp, forming a bipolar montage: F8-T4, F4-C4, Fz-Cz, F3-C3, F7-T3, P3-O1 and P4-O2. A discussion for the rationale of this configuration falls beyond the scope of this paper. It suffices to say that we are not looking for signals emanating from specific cortical sites; rather, the idea is to sense the EEG over a wide area of the cortex. The electrodes are plugged
into a biosignal amplifier and a real-time acquisition system. The analysis module performs power spectrum and Hjorth analyses in real-time. It generates two streams of control parameters. One stream contains information about the most prominent frequency band in the signal and is used to generate the music. The other stream contains information about the complexity of the signal and is used to control the tempo (beat) and dynamics (loudness) of the performance (Figure 3).

The present implementation activates generative rules for two different styles of music (e.g., Schumann-like and Beethoven-like), depending on whether the EEG is characterised by the presence of alpha rhythms or beta rhythms. Every time it has to produce a musical bar (or half-bar), it checks the power spectrum of the EEG at that moment and activates the generative rules accordingly. The system is initialised with reference tempo and amplitude values, which are modulated by the signal complexity analysis (Hjorth analysis).

5. CONCLUDING REMARKS

In order to have greater explicit (or intentional) control over the system, we are studying ways to develop methods to train subjects to achieve specific EEG patterns to control musical algorithms. We have initial evidence that this can be made possible using a technique known as biofeedback. Biofeedback is when biological information or signals are monitored electronically, which in turn feeds back information about your body or physiological state. This information is often displayed through audio-visual stimuli. As a result, the subject can learn to modify these signals and subsequently learn to gain greater control of the biological signals. Biofeedback technology has been used to treat and control a number of conditions; examples include migraine headaches and epilepsy. However, the problem of explicit objective control “by thinking” is a difficult one because while EEG analysis can help us to deduce whether a person is thinking, it is unlikely to give us any clear indication of what a person is actually thinking.

We acknowledge that the music produced by the system is of limited appeal for those interested in modern new music. Furthermore, the pieces produced by our computer-replication of musical style engine may not always sound convincing to discerning listeners. However, we decided to adopt the grammar-based approach inspired by the work of David Cope (2001) as a starting point for this research because such grammars are well understood and their use in music is well documented. Nevertheless, we are studying the possibility of using other interesting machine learning and generative techniques such as those proposed by Assayag and Dubnov (2004).

An aspect that needs to be addressed in future versions of our system is the non-ergonomic nature of the electrode technology for sensing the EEG. It can be uncomfortable and awkward to wear. There are various possibilities for innovations in the hardware design of EEG capture devices. Inexpensive auto-scanning and auto-negotiating wireless chips are now available and could be placed on the head along with the small preamplifiers. It is thus technically possible to build wearable EEG amplifiers with built-in signal processing and wireless data transmission.

Acknowledgements: The author would like to thank two former members of his post-graduate research team, Andrew Brouse and Bram Boskamp, for their valuable contributions to this project.

6. REFERENCES

E D Adrian and B H C Matthews (1934), The Berger Rhythm: Potential Changes from the Occipital Lobes in Man, Brain, 57, 4, pp. 355-85.

C Anderson and Z Sijercic (1996), Classification of EEG signals from four subjects during five mental tasks, Solving Engineering Problems with Neural Networks: Proceedings of the Conference on Engineering Applications in Neural Networks (EANN'96), London, UK.


